ЗБОРНИК РАДОВА

65. годишња конференција за електронику, телекомуникације, рачунарство, аутоматику и нуклеарну технику

ЕТРАН 2021

и

8. интернационална конференција за електротехнику, електронику и рачунарство

ИцЕТРАН 2021

PROCEEDINGS

8th International Conference on Electrical, Electronic and Computing Engineering

IcETRAN 2021

and

65th National Conference on Electronics, Telecommunication, Computing, Automatic Control and Nuclear Engineering

ETRAN 2021

Етно село Станишићи, Република Српска, 8 - 10. септембра 2021. године
Ethno Village Stanišići, , Republic of Srpska, 8 - 10, September, 2021

ISBN 978-86-7466-894-8

Издавачи
Друштво за ЕТРАН, Београд
Академска мисао, Београд

Published by
ETRAN Society, Belgrade,
Academic Mind, Belgrade

ЗБОРНИК РАДОВА

65. годишња конференција за електронику, телекомуникације, рачунарство, аутоматику и нуклеарну технику

ЕТРАН 2021

и

8. интернационална конференција за електротехнику, електронику и рачунарство

ИцЕТРАН 2021

PROCEEDINGS

8th International Conference on Electrical, Electronic and Computing Engineering

IcETRAN 2021

and

65th National Conference on Electronics, Telecommunication, Computing, Automatic Control and Nuclear Engineering

ETRAN 2021

Зборник радова - 65. Конференција за електронику, телекомуникације, рачунарство, аутоматику и нуклеарну технику,
Етно село Станишићи, 08-10.09.2021. године

Proceedings of Papers – 8th International Conference on Electrical, Electronic and Computing Engineering, IcETRAN 2021,
Ethno willage Stanišići, Republic of Srpska, Bosnia and Herzegovina

Главни уредник / Editor in Charge
Слободан Вукосавић / Slobodan Vukosavić

Издавачи / **Друштво за ЕТРАН, Београд и Академска мисао, Београд**
Published by **/ ETRAN Society, Belgrade, Academic Mind, Belgrade**

Израда / Production
Академска мисао, Београд / Academic Mind, Belgrade

Место и година издања / Place and year of publication
Београд, 2021. / Belgrade, 2021.

Тираж / Circulation
200 примерака / 200 copies

ISBN 978-86-7466-894-8

**ЕТРАН - Друштво за електронику, телекомуникације,
рачунарство, аутоматику и нуклеарну технику**

**ETRAN – Society for electronics, telecommunication,
computing, automatics and nuclear angineering**

Кнеза Милоша 9/IV, 11000 Београд / Kneza Milosa 9/IV, 11000 Belgrade

Телефон / Phone: 011 3233 957

Е-пошта / E-mail: office@etran.rs

www.etran.rs

ОРГАНИЗАТОРИ - ORGANIZERS

Друштво за ЕТРАН, Београд / ETRAN Society, Belgrade

**Универзитет у Бањој Луци - Електротехнички факултет /
University of Banja Luka - Faculty of Electrical Engineering,
Banja Luka, Republic of Srpska**

**Универзитет у Београду - Електротехнички факултет /
University of Belgrade - School of Electrical Engineering**

ПОДРШКА / SUPPORTED BY

IEEE – Institute of Electrical and Electronics Engineers, USA

Радови укључени у Зборник прихваћени су од стране рецензената и приказани на 65. годишњој конференцији Друштва за ЕТРАН (ЕТРАН 2021) и 8. Интернационалној конференцији (ИцЕТРАН 2021) које су одржане од 08. до 10. септембра 2021. године у Етно селу Станишићи, Република Српска.

Број пријављених радова за конференције ЕТРАН и ИцЕТРАН је 162. Рецензије радова обавило је укупно 266 рецензената. Просечан број рецензената по раду био је 2. Прихваћен је и на конференцији приказан 141 рад који су публиковани у овом зборнику.

Заједничка тематска сесија "Дигитална Србија и Република Српска" окупила је научнике, стручњаке, истраживаче, представнике високошколских установа и представнике државе који су изнели своје погледе на значај и развој информационих технологија и вештачке интелигенције, на њихову улогу у развоју привреде и на одговарајуће промене у образовном систему. Координатори сесије били су проф. др Бранко Докић и проф. др Мило Томашевић, док су активни учесници сесије били Мр Срђан Рајчевић, Министар за научнотехнолошки развој, високо образовање и информационо друштво у Влади Републике Српске, др Саша Стојановић, помоћник Министра за просвету, науку и технолошки развој Владе Србије, проф. др Мило Томашевић декан Електротехничког факултета Универзитета у Београду, проф. др Зоран Ђурић, декан Електротехничког факултета Универзитета у Бањој Луци и проф. др Божидар Поповић, декан Електротехничког факултета Универзитета у Источном Сарајеву.

Координатор специјалне седнице на секцији Метрологија, под насловом "Стохастичке методе у мерењима био је Владимир Вујичић. Координатор специјалне седнице на секцији Рачунарство и вештачка интелигенција, под насловом "Шта рачунари данас не могу" био је Бошко Николић. У оквиру секције за електроенергетику одржана је специјална седница "Електроенергетика у 21. веку" у организацији Одбора за енергетику САНУ.

Члан Председништва Предраг М. Петровић био је координатор заједничке тематске седнице организоване као омаж Милољубу Смиљанићу, Почасном члану Друштва за ЕТРАН и Генералном секретару Академије инжењерских наука. Уз поруку "Драги Мићо, дивљење и поштовање са захвалношћу" говорили су Предраг М. Петровић, Дејан Б. Поповић, Председник ЕТРАН-а.

Председник ЕТРАН-а, академик Дејан Б. Поповић био је координатор заједничке тематске седнице организоване као омаж академику Нинославу Стојадиновићу, бившем Председнику ЕТРАН-а, члану Председништва и Заслужном члану Друштва за ЕТРАН. Уз поруку "Остајемо да негујемо његове идеје" говорили су проф. др Данијел Данковић, Дејан Б. Поповић, Председник ЕТРАН-а и Братислав Миловановић, академик АИНС.

Посебно се захваљујемо организаторима из Републике Српске и домаћинима из Бијељине, који су допринели стварању услова за рад и плодну размену мишљења и критички осврт на резултате у оквиру свих секција.

Београд, 12.10.2021.
Академик Слободан Вукосавић
заменик председника ЕТРАН

The papers included in the Proceedings were selected in a peer review process and presented at the 65th annual conference of the ETRAN Society (ETRAN 2021) and at the 8th international Conference IcETRAN 2021, both held September 8 – 10, 2021 in Stanišići ethno-village, Republic of Srpska, Bosnia and Herzegovina.

The number of the submitted papers for the ETRAN and IcETRAN conferences was 162 in total. Peer reviewing was done by 266 reviewers. The average number of reviewers per paper was 2. A total of 141 papers was accepted, presented at the two conferences and published in full in these Proceedings.

The joint thematic session "Digital Serbia and Republic of Srpska" gathered scientists, experts, researchers, representatives of universities and governmental representatives who presented their opinions about the significance and development of information technologies and artificial intelligence, their role in the economic development and the corresponding changes in the educational system. Session coordinators were prof. dr. Branko Dokić and prof. dr. Milo Tomašević, while the active participants of the session were Mag. Sci Srdjan Rajčević, Minister of Scientific and Technological Development, Higher Education and Information Society in the government of the Republic of Srpska, dr. Saša Stojanović, Assistant Minister of Education, Science and Technological Development in the government of Republic of Serbia, prof. dr. Milo Tomašević, Dean of the School of Electrical Engineering, University of Belgrade, prof. dr. Zoran Djurić, Dean of the Faculty of Electrical Engineering, University of Banja Luka and prof. dr. Božidar Popović, Dean of the Faculty of Electrical Engineering, University of East Sarajevo.

The coordinator of the special session within the Metrology Section titled "Stochastic Methods in measurements" was Vladimir Vujičić. The coordinator of the special session within the Computers and Artificial Intelligence Section titled "What computers cannot do today" was Boško Nikolić. Within the Power Engineering Section a special session "Electric Power in 21st Century", organized by the Power Engineering Board of Serbian Academy of Sciences and Arts (SASA).

The member of the ETRAN Society Board Predrag M. Petrović was the coordinator of the plenary thematic session organized as a homage to late dr. Miloljub Smiljanić, Fellow of the ETRAN Society and Secretary General of the Serbian Academy of Engineering Sciences. With the message "Dear Mićo, admiration and respect with gratitude" the speakers were Predrag M. Petrović and academician Dejan B. Popović, ETRAN Society Chairman.

The ETRAN Society Chairman, academician Dejan B. Popović was the coordinator of the plenary thematic session organized as a homage to late academician Ninoslav Stojadinović, Member of ETRAN Society Board and a Fellow of ETRAN Society. With the message "We continue to forward his ideas" the speakers were prof. dr. Danijel Danković, academician Dejan B. Popović, ETRAN Society Chairman and Bratislav Milovanović, academician of Serbian Academy of Engineering Sciences.

We express our special gratitude to the organizers from the Republic of Srpska and our hosts from Bijeljina who contributed to creating working conditions and a fruitful interchange of opinions, as well as a critical review of the results within all sections.

Belgrade, October 12, 2021.
Academician Slobodan Vukosavić
Vice-Chairman of the ETRAN Society

Садржај/Contents

АКУСТИКА
/
ACOUSTICS

(AK/AKI)

Kriterijumi zvučnog komfora u prostorijama za vežbanje i izvođenje muzičkog programa

Dragana Šumarac Pavlović, Tatjana Miljković, Miloš Bjelić, Miomir Mijić

***Apstrakt*— Zvučni komfor u prostorima za vežbanje i izvođenje muzičkog programa definisan je kriterijumima iz tri domena: zvučne zaštite, kvaliteta zvučne slike i zaštite privatnosti. U ovim akustički osetljivim prostorima zvučni komfor se ne može regulisati opštim normativima iz više razloga. Radi se o prostorima u kojima se uobičajeno javljaju povišeni nivoi zvuka u prostorijama i istovremeno pooštreni kriterijumi za dozvoljene nivoe ambijentalne buke koja dospeva iz okruženja. U ovom radu dat je pregled različitih parametara koji se koriste u projektovanju ovakvih objekata, koji se oslanja na istraživanja, pre svega subjektivnih preferenci korisnika. Polazeći od spektralnih i dinamičkih karakteristika zvuka pojedinačnih muzičkih instrumenata i kriterijuma za zvučnu izolaciju i zaštitu privatnosti analizirani su mogući dometi zvučne zaštite prostora za vežbanje standardnim pregradnim konstrukcijama. Obzirom da je u planu izgradnja nove zgrade Fakulteta muzičke umetnosti u Beogradu, cilj ovog rada je pregled relevantnih standarda i principa koji se moraju poštovati u projektovanju da bi se u prostorima za muzičko obrazovanje ostvario zvučni komfor u svim njegovim apsektima.**

***Ključne reči*— probne sale, zaštita privatnosti, zvučna zaštita**

I. Uvod

Zvučni komfor, to jest kriterijumi akustičkog dizajna, zaštite privatnosti i zvučne izolacije u prostorijama za vežbu i muzička izvođenja u okviru muzičkih škola i fakulteta predstavljaju složen zadatak obzirom na brojne kriterijume koji treba da budu zadovoljeni. U ovim specifičnim objektima kriterijumi akustičkog dizajna prostorija i zvučne zaštite treba da zadovolje mnogobrojne specifične zahteve kako bi se postigla optimalna funkcionalnost prostora i realizovala podsticajna sredina za razvoj i muzičko izražavanje korisnika, bilo da se radi o prostorima za vežbu ili izvođenje. U ovakvim objektima uobičajeno se nalazi veliki broj prostorija različitih namena, kao što su male sale za probe pojedinačnih instrumenata, grupne sale za probe, velike sale za probe orkestra, studijski prostori i različiti prostori za muzička izvođenja (koncertne sale, operske sale, baletske sale). Osnovni nivo zahteva u projektovanju odnosi se na potreban volumen i dimenzije prostorija, optimalno vreme reverberacije i difuznost zvučnog polja u njima, kao i kontrolu sopstvenih modova prostorije. Sledeći nivo zahteva diktiran je karakteristikama pojedinačnih instrumenata kao što su različite spektralne karakteristike njihovog zvuka, dinamičke opsege pojedinih instrumenata i potrebu da se akustički dizajn optimizuje na osnovu kriterijuma optimalne glasnosti i podrške koju prostorija treba da pruži muzičarima kroz povratnu spregu sopstvenog kontrolnog mehanizma u izvođenju. U slučaju obrazovnih ustanova tema je i podrška koju prostor pruža u obrazovnom procesu. U domenu zvučne zaštite zahtevi se odnose na kriterijume za dozvoljene nivoe buke koja dospeva iz spoljašnje sredine, nivoa buke mašinskih sistema i potrebnu zvučnu izolovanost između prostorija.

Prostorije za vežbanje i muzička izvođenja su prostori u kojima je pažnja korisnika prostora usmerena prvenstveno na zvuk, to jest zvuk je predmet njihovog primarnog interesovanja i rada. Otuda svi kriterijumi za projektovanje treba da se oslanjaju s jedne strane na objektivne kriterijume, a s druge na subjektivna očekivanja korisnika. Korisnici ovakvih prostora, čak i kada imaju jasnu predstavu o tome šta im odgovara, ne mogu uvek svoje zahteve da izraze u korelaciji sa uobičajenim objektivnim parametrima.

Poslednjih 50 godina istraživači tragaju sa psihološkim mernim instrumentima kojima bi se opisao utisak o akustici prostorija i više od 100 godina traju napori da se predlože objektivne mere koje bi omogućile predviđanje perceptualnih karakteristika. Metodologijama koje su u upotrebi nema sasvim zadovoljavajućeg rešenja. [1,2].

U literaturi postoji generalni stav [3,4,5] da uslovi zvučnog komfora u postojećim prostorima za vežbu i izvođenje muzike ne odgovaraju sasvim potrebama korisnika. Projektanti često nisu upoznati sa kompleksnošću problema koji ne može biti jednostavno obuhvaćen zakonskim normama i projektantskim preporukama.

Zakonskom regulativom u našoj zemlji nisu obuhvaćeni ovakvi prostori u svoj svojoj složenosti [6]. U nacionalnom standardu SRPS U.J6.201 definisana je minimalna potrebna izolovanost između prostorija koje se graniče sa prostorijama za muzičko vaspitanje, i navodi se da je minimalan zahtev D_{wmin} = 60 dB. Za istu poziciju u zgradama za međuspratnu

Dragana Šumarac Pavlović – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar Kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: dsumarac@ etf.rs).

Tatjana Miljković – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar Kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: tm@ etf.rs).

Miloš Bjelić– Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar Kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: bjelic@ etf.rs).

Miomir Mijić– Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar Kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: emijic@ etf.rs).

konstrukciju se traže uslovi $R'_w \geq 60$dB i $L'_{wn} \leq 55$dB.

Zvučni komfor u muzičkim vežbaonicama bio je tema istraživanja u mnogim državama. Tako je u Norveškoj 2014. godine usvojen standard NS 8178 pod nazivom „*Acoustic criteria for rooms and spaces for music rehersal and performanse*“ [7]. U njemu su definisane potrebne dimenzije prostorija za muzička vežbanja i vreme reverberacije u njima za različite tipove muzike i različit broj članova muzičke grupe. Razmatrane su tri kategorije muzičkog sadržaja: prirodni instrumenti, muzički instrumenti male snage i muzika koja koristi elektroakustički sistem za ozvučavanje. U Velikoj Britaniji publikovan je vodič za projektovanje škola u kome je jedno poglavlje posvećeno preporukama za dizajn prostora namenjenih muzičkom obrazovanju [4,8].

U ovom radu su kao polazna osnova za utvrđivanje objektivnih kriterijuma poslužili rezultati jednog velikog istraživanja koje je sprovedeno na Škotskoj kraljevskoj akademiji za muziku i dramu (*Royal Scottish Academy od Music and Drama*) [5]. Osnovna ideja ovog istraživanja bila je da se na osnovu subjektivnih stavova korisnika prostora utvrdi koji su najvažniji faktori u dizajnu prostora za muzičko vežbanje, s obzirom da ih studenti aktivno koriste više od 42 sata nedeljno. Cilj istraživanja je bio da se utvrdi optimalno vreme reverberacije, tolerantnost na ambijentalnu buku i fizičke karakteristike prostora u korelaciji sa akustičkim karakteristikama. Istraživanjem je obuhvaćeno 22% ukupnog broja studenata muzike, a izabrani su oni koji sviraju drvene duvačke instrumente, limene duvačke, žičane instrumente i pevači. Istraživanje je sprovedeno u više faza u okviru kojih je namenski napravljeno šest različitih vežbaonica sa velikim razlikama u dimenzijama, akustičkim karakteristikama i različitim nivoima ambijentalne buke.

Rezultati analize su pokazali da je za 86% anketiranih studenata primetna saobraćajna buka, ali da većina smatra da je glavni ometajući faktor buka koja potiče iz susednih vežbaonica i to prevashodno kada je u pitanju isti instrument. Utvrđeno je da su prihvatljivi nivoi buke:

- NC30 za buku ventilacionih sistema
- NC25 za buku saobraćaja
- NC15 za buku od istog instrumenta iz susednih vežbaonica.

Pri tome, preko 90% ispitanika se slaže da je nivo od 45 dB(A) koji potiče od saobraćajne buke manje uznemirujući od nivoa 23 dB(A) koji potiče iz susedne vežbaonice.

Prema mišljenju ispitanika optimalno vreme reverberacije prostorija za muzička vežbanja je oko 0.7 s, pri čemu je poželjno da postoje uslovi za promenljivu akustiku kojom bi se vreme reverberacije moglo menjati u opsegu od 0.5 do 0.9s.

Najzad, utvrđeno je da su optimalne dimenzije prostorija za individualno vežbanje 15-20 m^2.

U britanskom dokumentu BB93 definisane su kriterijumi za nivo ambijentalne buke koja potiče iz susedne prostorije za nekoliko različitih slučajeva [8]. Oni su prikazani u Tabeli I. Uzimajući u obzir nivo aktivnosti u pobudnoj prostoriji i stepen tolerancije na nivo ambijentalne buke u tabeli II su definisane minimalne vrednosti izolovanosti između prostorija za različite kombinacije.

TABELA I
KLASIFIKACIJA PROSTORIJA PREMA TOME KOLIKI SE NIVO ZVUKA U NJIMA GENERIŠE, KOLIKO SU OSETLJIVE NA SPOLJAŠNJU BUKU I KRITERIJUM DOZVOLJENOG NIVOA BUKE [8]

Tip prostorije	Nivo na pobudnoj strani	Tolerancija na prijemnoj strani	Kriterijum dB(A)
Učionica	Veoma visok	Niska	**35**
Mala vežbaonica	Veoma visok	Niska	**35**
Grupna vežbaonica	Veoma visok	Veoma niska	**30**
Koncertna slala	Veoma visok	Veoma niska	**30**
Studio za snimanje	Veoma visok	Veoma niska	**30**
Kontrolna soba	Visok	Niska	**35**

U norveškom standardu NS 8178 polazna osnova za utvrđivanje optimalnog vremena reverberacije za zadati volumen prostora su dva podatka: prosečna jačina zvuka *G* (*stength*) i tipične vrednosti nivoa zvučne snage koju generišu pojedini instrumenti kada se svira dinamikom *forte*.

Polazeći od ukupnog optimalnog nivoa zvuka i spektralnih karakteristika pojedinih grupa instrumenata moguće je utvrditi domete izolacione moći standardnih pregradnih konstrukcija kako bi se zadovoljili uslovi izolovanosti između susednih prostora za muzičku vežbu.

Za nivo udarnog zvuka u svim prostorijama u kojima se izvodi ili snima muzički program definisan je kriterijum $L'_{nT,w} \leq 55$ dB

TABELA II
MINIMALNE VREDNOSTI MERODAVNE IZOLOVANOSTI ZA RAZLIČITE KOMBINACIJE AKTIVNOSTI I TOLERANCIJE NA NIVO AMBIJENTALNE BUKE U RAZLIČITIM PROSTORIMA [8]

Min *DnT* (dB)		Nivo aktivnosti u pobudnoj prostoriji			
Tolerancija na nivo amb. buke	Visoka	30	35	45	55
	Srednja	35	40	50	55
	Niska	40	45	55	55
	Veoma niska	45	50	55	60

II. NIVO ZVUKA I DINAMIČKI OPSEG

Nivo zvuka u prostoriji koji generišu prirodni instrumenti zavisi od više faktora:

- tipa i broja instrumenata,
- načina sviranja, dinamike
- zapremine prostorije
- vremena reverberacije u prostoriji.

Muzički instrumenti se prema načinu generisanja tona dele na žičane instrumente sa gudalom i okidanjem žica, drvene duvačke, limene duvačke i udaraljke. Način generisanja tona određuje nivo zvuka koji proizvode kao i spektralni sadržaj. Nivo zvuka zavisi takođe i od dinamike sviranja. Kao ilustracija, u tabeli III pobrojane su dinamike sviranja i prosečni nivoi koji se na taj način postižu.

Vrednosti prikazane u tabeli odgovaraju vrednostima nivoa zvuka simfonijskog orkestra koji se očekuju u koncertnim salama. Iz tabele možemo da zaključimo da je dinamički opseg koji proizvodi simfonijski orkestar u toku kompozicije može dostići nivo od 60 dB. Zbog velikog dinamičkog opsega uobičajeno je da se u proračunima razmatraju nivoi koji odgovaraju dinamici *forte*. Majer je u svojim istraživanjima pokazao da se optimalni nivoi za slušaoce kreću u rasponu 85-92 dB, pa su te vrednosti uzete u standardima kao osnova za proračun svih vrednosti vezanih za zvučni komfor.

TABELA III
RAZLIČITE DINAMIKE SVIRANJA I PROSEČNI NIVOI ZVUKA KOJI SE POSTIŽU U KONCERTNIM SALAMA [10]

Oznaka dinamike	Značenje	Deskripcija	Približan nivo SPL
ppp	*piano pianissimo*	ekstremno tiho	**45-50 dB**
pp	*pianissimo*	jako tiho	**55-60 dB**
p	*piano*	tiho	**65-70 dB**
mf	*mezzo forte*	srednje glasno	**75-80 dB**
f	*forte*	glasno	**85-90 dB**
ff	*fortissimo*	jako glasno	**95-100 dB**
fff	*forte fortissimo*	ekstremno glasno	**105-110 dB**

TABELA IV
PODACI O ZVUKU RAZLIČITIH INSTRUMENATA: NIVOI ZVUČNE SNAGE PRI DINAMICI *pp*, *f* i *f*, DINAMIČKI OPSEG D I *k* FAKTOR [4]

Tip instrumenta	$L_{w(pp)}$ dB	$L_{w(ff)}$ dB	*D* dB	$L_{w(f)}$ dB	*k*
Violina	*65*	97	32	89	**0,8**
Viola	*68*	93	25	87	**0,5**
Violončelo	*67*	97	30	90	**1,0**
Kontrabas	*75*	97	22	92	**1,6**
Flauta	*77*	96	19	91	**1,3**
Klarinet	*74*	101	27	93	**2,0**
Saksofon	*87*	101	14	98	**6,3**
truba	*87*	106	19	101	**12,6**
trombon	*89*	109	20	104	**25,1**

U tabeli IV prikazani su osnovni podaci o. nivoima zvučne snage i dinamičkim opsezima koje dostižu pojedinačni instrumenti sa različitim dinamikama sviranja [4].

III. PRORAČUN NIVOA ZVUKA U RAZLIČITIM PROSTORIJAMA

Kada je poznat nivo zvučne snage izvora zvuka nivo zvučnog pritiska u prostoriji se može proceniti na osnovu izraza:

$$L_p = L_w + 10\log\left(\frac{1}{4\pi r^2}\right) + 10\log\left(\frac{4(1-\bar{\alpha})}{S\bar{\alpha}}\right)(dB) \qquad (1)$$

Kritično rastojanje u prostorijama u funkciji zapremine i vremena reverberacije prikazano je na slici 1. Vidi se da u prostorijama za vežbanje ono ne prelazi vrednost 0,7 m. To znači da se samo muzičar koji svira nalazi u zoni direktnog zvuka njegovog instrumenta, dok se svi ostali prisutni u prostoriji nalaze van zone direktnog zvuka. Na osnovu toga proizilazi da se nivo zvuka koji slušaju ostali prisutni može približno odrediti izrazom:

$$L_p = L_w + 10\log\left(\frac{4(1-\bar{\alpha})}{S\bar{\alpha}}\right)(dB) \qquad (2)$$

Pojačanje koje prostorija unosi opisuje se veličinom G koja se naziva jačina (*strength*). Ona je:

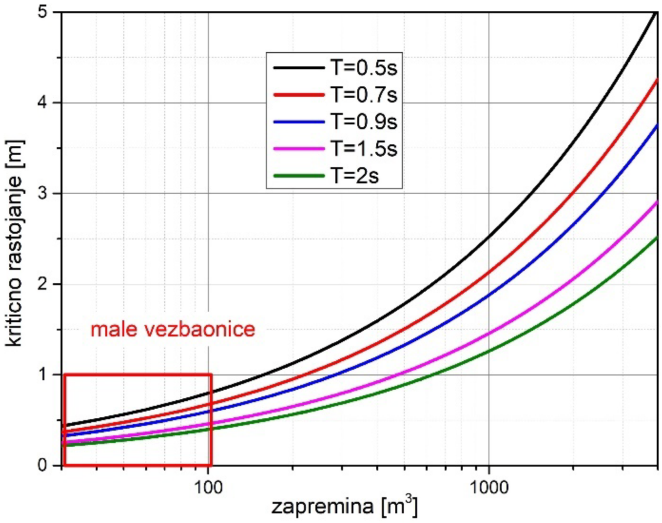
$$G = L_p - L_{pdir}(r_0 = 10m)(dB) \qquad (3)$$

G se može proceniti na osnovu srednjeg koeficijenta apsorpcije i ukupne unutrašnje površine u prostoriji na sledeći način:

$$G = 10\log\left(\frac{4(1-\bar{\alpha})}{\bar{\alpha}S}\right) - 10\log\left(\frac{1}{4\pi r_0^2}\right) \cong 31 + 10\log\left(\frac{4(1-\bar{\alpha})}{\bar{\alpha}S}\right)(dB) \qquad (4)$$

Za svaku geometrijsku formu prostorije može se definisati parametar koji je nazvan "faktor oblika", i koji predstavlja jedan mogući kvantifikator njene geometrijske forme [9]

$$k = \frac{\sqrt[3]{V}}{\sqrt[2]{S}}, \quad S = \frac{V^{2/3}}{k^2}(m^2), \quad S = \frac{0.16V}{\bar{\alpha}T}, \quad \bar{\alpha} = \frac{0.16V^{1/3}}{k^2 T} \qquad (5)$$



Slika 1 Kritično rastojanje u prostorijama u funkciji zapremine za različite vrednosti vremena reverberacije

Za jednu formu prostora konstantan je faktor oblika nezavisno od stvarnih dimenzija prostorije. Faktor oblika za paralelopipedne prostorije različitih proporcija kreće se u rasponu 0,35-0,4 [9]. Opseg vrednosti srednjeg koeficijenta apsorpcije u prostorima za vežbu označen je na slici 2.

Nivo zvuka u prostoriji može se izraziti preko vrednosti G i akustičke snage izvora [2]. Uobičajeno je da se nivo zvuka instrumenata razmatra pri sviranju sa dinamikom *forte*.

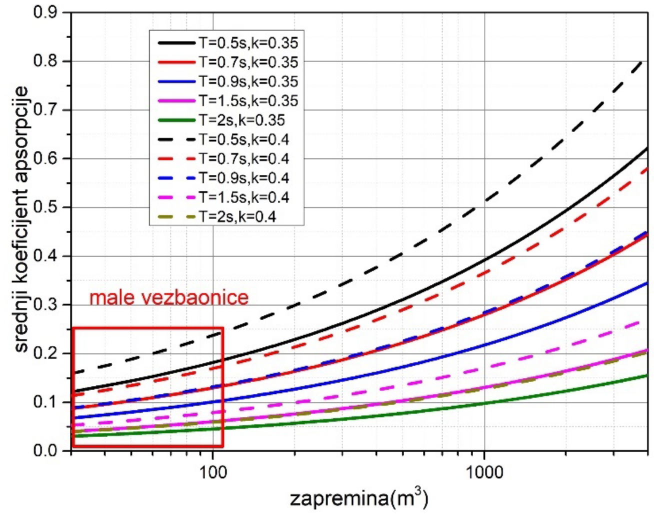
$$L_p(f) = L_w(f) + G - 31(dB) \qquad (6)$$

Ukoliko u prostoriji postoji više izvora ukupan nivo zvuka može se proceniti pomoću izraza [11]:

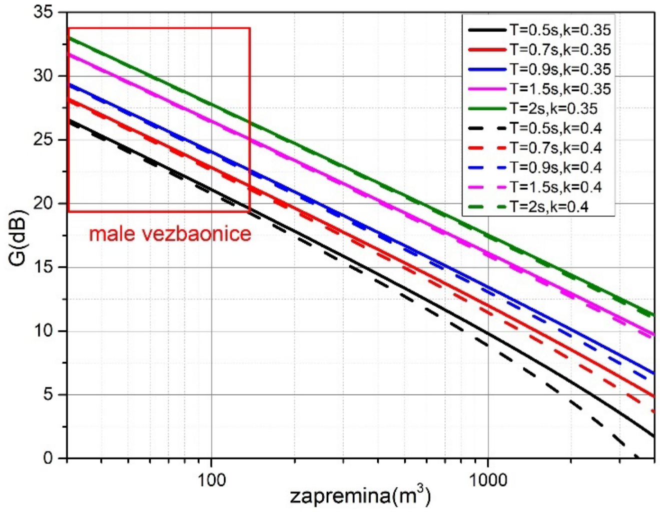
$$L_p(f) = G + 59 + 10\log\sum_i n_i k_i (dB) \qquad (7)$$

gde je n_i broj instrumenata iz iste grupe i k_i faktor definisan u tabeli IV.

Jačina zvuka u prostoriji menja se zavisno od njenog volumena, vremena revreberacije i faktora oblika. Na slici 3 prikazane su promene G u funkciji zapremine za različite vrednosti vremena reverberacije i faktora oblika. U vežbaonicama manjih zapremina za raspone vremena reverberacije od 0,5 s do 0,7 s jačina zvuka G ima vrednosti u opsegu 20-30 dB.



Slika 2 Srednji koeficjient apsorpcije u prostorijama u funkciji zapremine za različite vrednosti vremena reverberacije i različite granične vrednosti faktora oblika



Slika 3 Vrednosti parametra G u funkciji zapremine, vremena reverberacije i faktora oblika; označen je opseg vrednosti u kome se nalaze male muzičke vežbaonice.

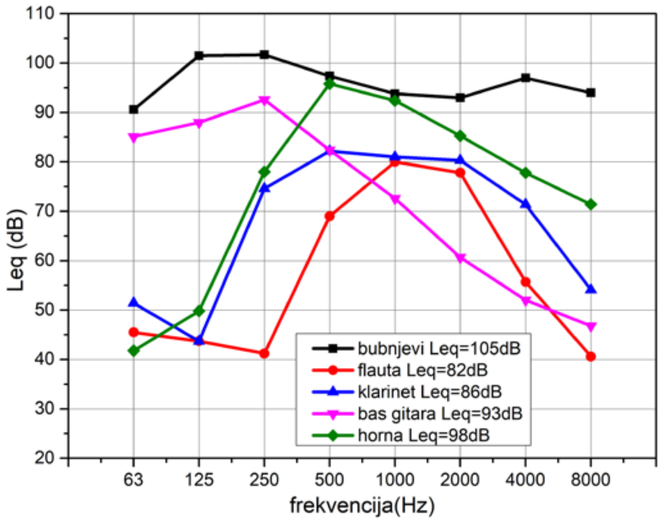
IV. Spektralne karakteristike instrumenata

Različiti instrumenti osim što se mogu okarakterisati različitim nivoima zvučne snage i različitim dinamičkim opsegom, značajno se međusobno razlikuju po opsegu frekvencija u kojima dominanto stvaraju zvuk, pa prema tome i po spektralnim karakteristikama. Pored ukupnog nivoa zvuka u prostoriji spektralna karakteristika može značajno da utiče da princip akustičkog dizajna i zvučne zaštite prostora.

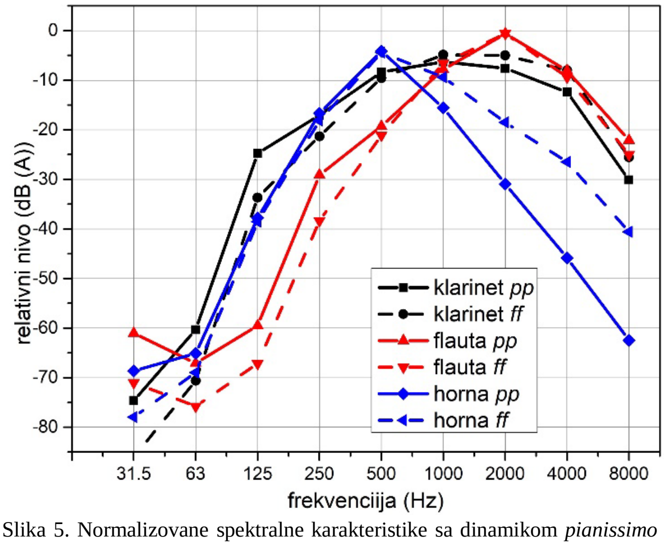
Takođe, spektralna karakteristika se menja u zavisnosti od dinamike sviranja. Mayer je utvrdio da je povećanjem dinamike signala menja nagib spektralne karakteristike tako što se dominanto povećava energija na višim harmonicima [11]. To dovodi i do toga da dinamika koji izvođač čuje u zoni direktnog zvuka instrumenta nije ista kao dinamika koju percipiraju slušaoci na udaljenim mestima u sali, pogotovo kada je prisutno značajno smanjenje vremena reverberacije sa frekvencijom.

Na slici 4 prikazane su izmerene vrednosti ekvivalentnog nivoa zvuka u manjim vežbaonicama različitih instrumenata [10]. Ovi podaci mogu se koristiti kao relevantni za procenu nivoa zvučne snage ovih izvora. Razmatrano je nekoliko kategorija instrumenata i to baterija bubnjeva, flauta, klarinet, bas gitara i francuski rog.

Na slici 5 prikazani su normalizovani spektri flaute, klarineta i horne sa dinamikom sviranja *pianissimo* i *fortissimo.* Spektralne karakteristike su ponderisane A filtrom. Kod svih instrumenata uočljivo je manje ili više izdizanje spektra na visokim frekvencijama, što instrumentalisti daje subjektivni doživljaj veće dinamike koju postiže sviranjem.



Slika 4. Spektralne karakteristike i ukupni ekvivalentni nivo za različite instrumente.



Slika 5. Normalizovane spektralne karakteristike sa dinamikom *pianissimo* i *fortissimo* ponderisane A filtrom

V. Potrebna izolaciona moć pregradnih konstrukcija

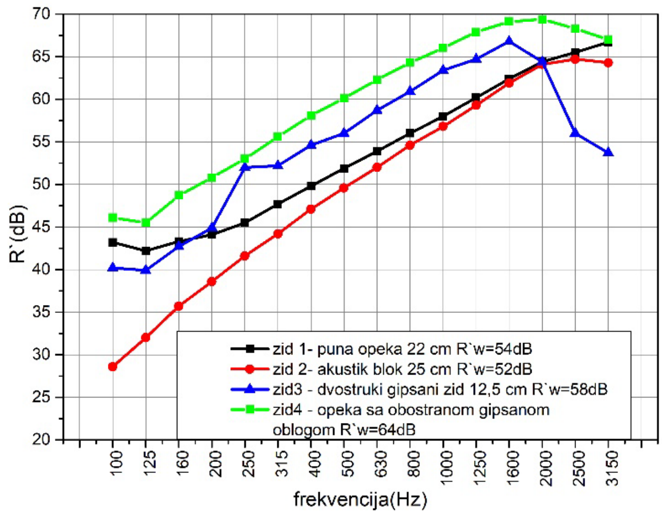
Istraživanja subjektivnog doživljaja muzičara kada vežbaju u malim i srednjim vežbaonicama pokazala su da su u pogledu zvučne izolacije i zaštite privatnosti u objektu kritične pozicije one koje se neposredno graniče sa prostorijama u kojima se pojavljuju isti instrumenti [5]. Kod muzičara koji sviraju određeni instrument dodatno je pojačano analitičko slušanje i osetljivost na za njih prepoznatljiv zvuk.

U literaturi je ranije konstatovano da su najstrožiji kriterijumi za zvučnu izolovanost između dva susedna prostora u slučaju da u njima vežbaju muzičari koji sviraju isti instrument. U tom slučaju nivo zvuka koji prolazi u susednu prostoriju mora zadovoljavati kriterijum NR15, odnosno nivo zvuka 23 dB(A) [5].

Polazeći od ovih zaključaka analizirane su izolovanosti koje se postižu sa nekoliko standardnih pregradnih konstrukcija:

- ZID 1 - zid od pune opeke (22cm) dvostruko malterisan;
- ZID 2 – zid od blokova debljine 25 cm obostrano malterisan;
- ZID 3 – sendvič zid od dvostrukih gipsanih obloga;
- ZID 4 – zid od pune opeke sa dvostranom gipsanom oblogom od dve ploče gipsa.

Građevinske izolacione moći razmatranih pregradnih konstrukcija prikazane su na slici 6.



Slika 6 . Građevinske izolacione moći za četiri tipa pregrada

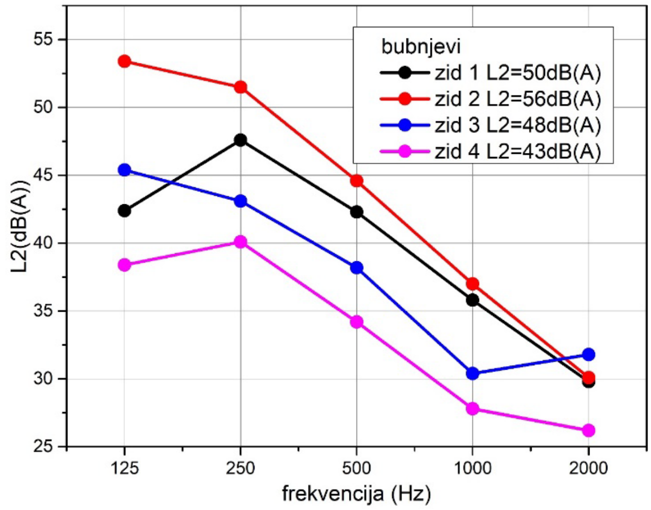
Za proračune izolovanosti posmatran je slučaj malih vežbaonica veličine 40 m^3 koje deli pregradni zid površine 10 m^2 i u kojima je srednji koeficijent apsorpcije $\alpha = 0{,}15$. Izolovanost D između njih izračunava se na osnovu vrednosti građevinske izolacione moći pregrade koja ih deli, površine te pregrade S i ukupne apsorpcije u prijemnoj prostoriji:

$$D = R' - 10\log\frac{S}{A_{prijem}}$$

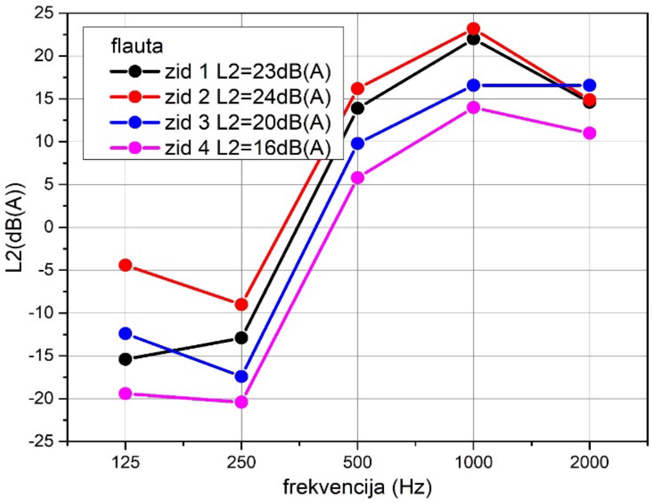
Za usvojene dimenzije i akustičke karakteristike vežbaonice izračunati su nivoi u prijemnoj prostoriji kada su u predajnoj prostoriji usvojeni ekvivalentni nivoi prikazani na slici 5. Za svaki od analiziranih instrumenata prikazane su spektralne karakteristike buke u prijemnoj prostoriji koja nastaje preslušavanjem (slike 7,8,9,10 i 11)

Analizom nivoa zvuka u prijemnoj prostoriji pri pobudama različitim instrumentima pokazano je da standardne pregradne konstrukcije nisu u mogućnosti da zadovolje stroge kriterijume zvučne izolovanosti. U slučaju baterije bubnjeva ni sa obostranim oblaganjem dvostrukim gipsanim oblogama na potkonstrukciji od mineralne vune nije moguće postići željeni nivo izolovanosti. Šta više, nivoi u prijemnoj prostoriji premašuju i nivoe dozvoljene buke koja potiče od saobraćaja i sistema za ventilaciju. To su situacije u kojima se zadovoljavajuća izolovanost može postići samo uvođenjem tampon zona oko vežbaonica za bučne instrumente.

Prikazani rezultati pokazuju da u slučaju flaute i klarineta, zbog njihove znatno manje zvučne snage i drugačijeg oblika spektra zvuka, moguće je postići traženu izolovanost i sa standardnim zidovima, to jest bez dodatih gipsanih obloga. U slučaju bas gitare i francuskog roga standardnim pregradama građevinskim nije moguće postići adekvatnu zvučnu izolaciju između prostorija za vežbanje.



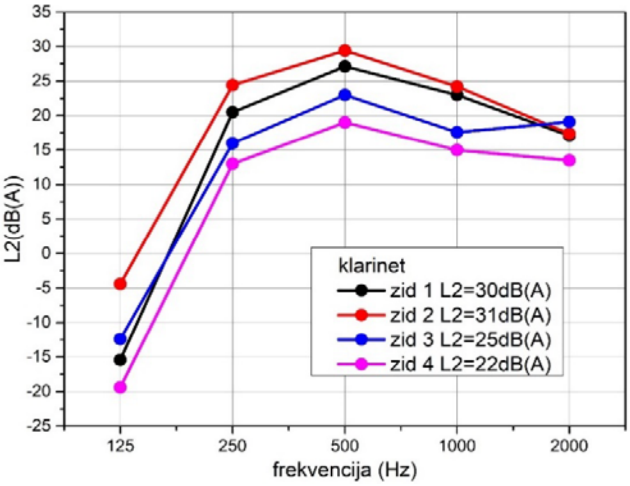
Slika 7. Spektralne karakteristike nivoa zvuka u prijemnoj prostoriji kada je izvor zvuka baterija bubnjeva



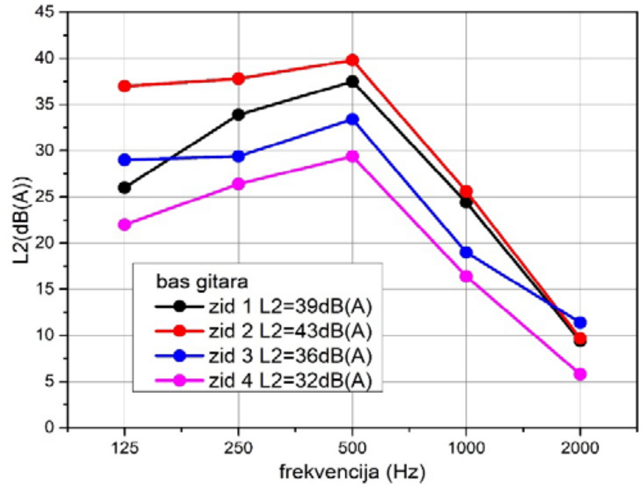
Slika 8. Spektralne karakteristike nivoa zvuka u prijemnoj prostoriji kada je izvor zvuka flauta

VI. ZAKLJUČAK

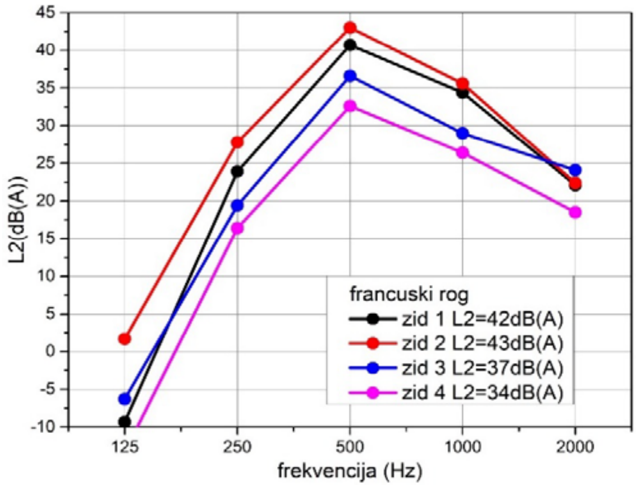
U radu je prikazan pregled poznatih metodologija i standarda koji se odnose na zvučni komfor u zgradama sa prostorijama za vežbanje i izvođenje muzičkog programa.



Slika 9. Spektralne karakteristike nivoa zvuka u prijemnoj prostoriji kada je izvor zvuka klarinet



Slika 10. Spektralne karakteristike nivoa zvuka u prijemnoj prostoriji kada je izvor zvuka bas gitara



Slika 11. Spektralne karakteristike nivoa zvuka u prijemnoj prostoriji kada je izvor zvuka bas gitara i sa različitim pregradama između njih

Prikazan je pregled nekih rezultata koji su dobijeni u subjektivnim istraživanjima vezanim za očekivanja muzičara u pogledu zvučnog komfora i definisane su osnovne veličine od kojih se polazi u procesu akustičkog dizajna i projektovanja zvučne zaštite u takvim objektima. Analizirana je zvučna izolovanost i privatnost koja se postiže sa četiri standardne pregrade za slučaj različitih instrumenata kao izvora zvuka.

Analiza rezultata pokazala je da se pri izboru pregradnih konstrukcija mora voditi računa o realnim nivoima zvuka i spektralnim karakteristikama instrumenata kako bi se pronašla adekvatna rešenja koja obezbeđuju traženu međusobnu izolovanost prostorija. Ta rešenja moraju u određenim slučajevima da podrazumevaju tampon zone između prostorija.

LITERATURA

[1] Weinzierl, S., Vorlander M., „*Room Acoustical Parameters as Predictors of Room Acoustical Impression: What Do We Know and What Would We Like to Know?*“, Acoustic Australia, DOI 10.1007/s40857-015-0007-6

[2] SRPS ISO 3382-1:2009, Acoustics - Measurement of room acoustic parameters - Part 1: Performance spaces.

[3] Knofel B., Weisheit L., Troge, J, “ *Musicians and their practice rooms: What do they think about present room acoustics and would they prefer?”,* Conference: Euronoise 2018, 11th European Congress and Exposition on Noise Control Engineering : Reduce Noise to Improve Life

[4] Rindel, J. H.,”*Rooms for music – Acoustical needs and requirements*”, BNAM2014

[5] Lamberty,D.C.,(1980) “*Music Practice Rooms”, Journal od Sound and Vibration*, Vol 60,,No1, p149-155

[6] SRPS ISO UJ.6.201, “ Akustika u zgradarstvu- Tehnički uslovi za projektovanje i građenje zgrada”, 1989

[7] NS 8178:2014, “*Acoustic criteria for rooms and spaces for music rehearsal and performance”.* (In Norwegian).Standard Norge, Oslo, 2014

[8] BB93: ”*Acoustic design of School, performance standard*”, Education Funding Agency, 2015

[9] D. Šumarac-Pavlovic´ and M. Mijic, ”*An approach to numerical quantification of room shape and its function in diffuse sound field model (L),* J. Acoust. Soc. Am. 140 (4), October 2016

[10] James A., Thompson A., Rees I.,”*School Music Rooms- designing Beyond BB93*, Proceedings of the Institute of Acoustics, 2005

[11] J. Meyer, “*Acoustics and the performance of music*”, Springer, 2009

ABSTRACT

Sound comfort in the spaces for rehearsing and performing music program is defined by criteria from three domains: sound protection, sound image quality and privacy protection. The problem is especially significant in the buildings of music schools and faculties. In such acoustically sensitive spaces, sound insulation cannot be regulated by general legislation for two reasons. First, these are spaces where high sound levels usually occur compared to standard spaces such as those in other buildings. Second, they have strict criteria for the allowed levels of ambient noise coming from the environment. That is reflected in the special measures of sound insulation, as well as the control of all systems in the building. Starting from the spectral and dynamic characteristics of the musical instruments sound and the criteria for sound insulation and privacy protection, the sound insulation ranges of standard partition structures were analyzed. Special attention is paid to individual and group rehearsal spaces. Given that the construction of a new building for the Music Academy in Belgrade is planned, the aim of this paper is to review all relevant standards and principles that must be respected in order to achieve sound comfort in music education spaces in all its aspects.

Criteria for acoustic comfort in the premises for practicing and performing music program

Dragana Šumarac Pavlović, Tatjana Miljković, Miloš Bjelić, Miomir Mijić

Uticaj estimacije frekvencija harmonika na procenu koeficijenta inharmoničnosti čembala

Tatjana Miljković, Jovana Damnjanović, Jelena Ćertić, Dragana Šumarac Pavlović

***Apstrakt*—Inharmoničnost je pojava koja se javlja kod muzičkih instrumenata koji se teorijski smatraju harmoničnim, i predstavlja odstupanje frekvencija parcijala tona od celobrojnih umnožaka osnovne frekvencije tona. Za žičane muzičke instrumente, u literaturi se definiše koeficijent inharmoničnosti, kao mera odstupanja od idealne harmoničnosti. U prethodnom istraživačkom radu pokazano je da automatski algoritam za procenu koeficijenta inharmoničnosti tonova klavira za pojedine tonove iz registra ne vrši dobru estimaciju koeficijenta inharmoničnosti. U ovom radu razmatran je uticaj tačnosti procene frekvencija harmonika na procenu koeficijenta inharmoničnosti. Sprovedena je uporedna analiza dve metode za procenu spektra signala, i to procena spektra na osnovu autoregresivnog AR modela i metodom diskretne Furijeove transformacije DFT. Testiranje predloženih metoda za procenu spektra izvršeno je na realnim tonovima čembala. Ustanovljeno je,na osnovu trenda koeficijenta inharmoničnosti računatog pomoću obe metode za procenu spektra, da je metoda AR modelom superiornija, odnosno da se dobijaju relevantni rezultati za koeficijent inharmoničnosti na celom opsegu tonova od interesa.**

***Ključne reči*—Inharmoničnost, žičani instrumenti, čembalo, procena spektra.**

I. Uvod

„Idealna" žica predstavlja žicu zategnutu na oba kraja, koja se ne odlikuje krutošću. Jedna od glavnih karakteristika takve žice, sa aspekta stvaranja zvuka, je što pri njenom okidanju dolazi do formiranja niza harmonika koji predstavljaju celobrojne umnoške osnovne frekvencije žice. U realnosti nemoguće je posmatrati žicu kao element sa svojstvima idealne fleksibilnosti, već je neophodno uvesti dodatne parametre poput krutosti, debljine, težine i napetosti žice. Krajem XIX veka *John William Strutt*, lord *Rayleigh*, razmatrajući oscilovanje žica klavira, ustanovio je da krutost same žice utiče na parcijale tonova, tako što frekvencije parcijala tonova na klaviru odstupaju od predefinisanih frekvencija parcijala idealne žice [1]. Pojava odstupanja frekvencije parcijala tona od celobrojnih umnožaka osnovne frekvencije tona naziva se inharmoničnost.

Matematičko-fizički modeli [2], [3], [4] koji u osnovi opisuju inharmoničnost kod klavira, definišu principski jednostavne relacije koje pojavu opisuju preko koeficijenta inharmoničnosti, koji se definiše za svaku žicu instrumenta.

Postoje različite metode i algoritmi za procenu koeficijenta inharmoničnosti na osnovu tonova instrumenta od interesa. *Galembo* i *Askenfelt* [5] predložili su jednu od metoda za procenu koeficijenta inharmoničnosti koja se zasniva na inharmoničnom komb filtru, realizovanom u frekvencijskom domenu. Takođe, tehnike poput određivanja visine tona, kepstralne analize i HPS (*Harmonic Product Spectrum*) su korišćene za estimaciju inharmoničnosti [6], [7]. Jedan od algoritama koji se odlikuje automatskim postupkom za estimaciju koeficijenta inharmoničnosti jeste PFD (*Partial Frequency Deviation*) algoritam [8]. Princip rada algortima zasniva se na minimiziranju devijacije, koja je nastala kao razlika frekvencija parcijala procenjenih iz spektra tona i frekvencija parcijala računatih na osnovu matematičkog modela. Sam proces minimizacije frekvencijske devijacije se odvija iterativnim postupkom, gde je krajnji produkt algoritma estimirana vrednost koeficijenta inharmoničnosti tona.

U prethodnom istraživačkom periodu analizirani su različiti klaviri na osnovu PFD algoritma za procenu koeficijenta inharmoničnosti. Izvršena je analiza keoficijenta inharmoničnosti procenjena sa različitim brojem parcijala tonova. Takođe, utvrđeno je da se takav algoritam može koristiti za procenu koeficijenta inharmoničnosti tonova klavira za skoro celokupni registar instrumenta [9]. Za pojedine tonove iz opsega klavira, iako je došlo do konvergencije samog algoritma, estimirana vrednost koeficijenta inharmoničnosti nije bila tačna. Razlog za lošu procenu koeficijenta inharmoničnosti pojedinih tonova leži u lošoj proceni frekvencija parcijala tona iz spektra signala.

U ovom radu razmatran je uticaj tačnosti procene frekvencija harmonika na procene koeficijenta inharmoničnosti.. Korišćen je pomenuti iterativni algoritam PFD, pri čemu je akcenat stavljen na prvi korak adaptivnog algoritma, odnosno procenu frekvencija tona iz spektra. Izvršena je uporedna analiza dve metode za procenu spektra signala i to procena spektra na osnovu AR (*Autoregressive*) modela i metodom DFT (*Discrete Fourier transform*). Uticaj tačnosti estimacije frekvencija harmonika na procenu koeficijenta inharmoničnosti testiran je na tonovima čembala. Za potrebe rada izvršeno je snimanje tonova čembala na Fakultetu muzičke umetnosti u Beogradu. Čembalo je izabrano kao instrument za analizu, jer se spektri njegovih tonova odlikuju velikim brojem harmonika. Prisutnost velikog

Tatjana Miljković – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar Kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: tm@ etf.bg.ac.rs).

Jovana Damnjanović – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar Kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: dj193200m@etf.bg.ac.rs).

Jelena Ćertić – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar Kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: certic@ etf.bg.ac.rs).

Dragana Šumarac Pavlović – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar Kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: dsumarac@etf.bg.ac.rs).

broja parcijala u spektru tona je značajna, zbog činjenice da su fizičke pojave koje opisuju fenomen inharmoničnosti uočljivije na višim parcijalima tona.

Rad je oragnizovan kako sledi. U drugom poglavlju date su osnovne razlike između čembala i klavira, poput mehanizma stvaranja zvuka i spektralne odlike tona ovih instrumenata. U narednom poglavlju prikazane su metodologije za procenu spektra korišćene u iterativnom algoritmu za procenu koeficijenta inharmoničnosti. U četvrtom poglavlju prikazani su eksperimentalni rezultati i diskusija dobijenih rezultata. Na kraju izneti su zaključci do kojih se došlo u ovom istraživanju.

II. OSNOVNE RAZLIKE IZMEĐU ČEMBALA I KLAVIRA

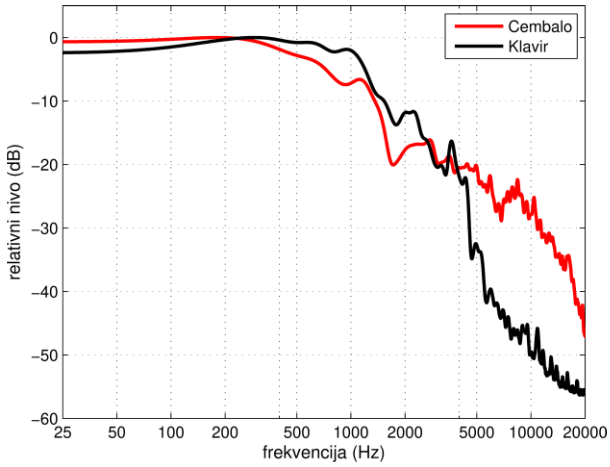
Žičani instrumenti su svi oni instrumenti kod kojih se zvuk stvara vibriranjem žice. S obzirom na različite načine pobuđivanja žice na vibriranje žičani instrumenti se mogu klasifikovati u 3 podgrupe, i to na: gudačke instrumente (violina, violončelo), trzane (gitara, čembalo) i udarne žičane instrumente (klavir) [10].

Čembalo je predstavnik žičanih instrumenata, koji spadaju u podgrupu trzanih žičanih instrumenata sa klavijaturom. Iako izgledom podseća na klavir, postoji nekoliko suštinskih razlika između ova dva instrumenta. Jedna od razlika je u materijalu od kojih su sačinjena ova dva instrumenta. Čembalo je izgrađeno od laganog drvenog okvira, tankih žica i nema pedale. Dok je klavir izgrađen od teškog gvozdenog okvira, debljih žica gde su neke od njih i obmotane i pedala. Takođe, čembalo može imati jednu, dve ili tri klavijature, koje se nazivaju manueli. Opseg tonova koji se može odsvirati na čembalu je manji u odnosu na klavir i uglavnom iznosi 5 oktava, dok opseg klavira uglavnom iznosi sedam oktava [11].

Značajna razlika između ova dva instrumenta je u mehanizmu nastanka zvuka. Kod klavira zvuk nastaje udarcem čekića o veoma zategnutu žicu, dok kod čembala zvuk nastaje okidanjem blago rastegnute žice perom. Pero koje služi za okidanje žice nalazi se na pokretnom jezičku, smeštenom na vrhu stubića, koji svojim donjim krajem naleže na zadnji kraj dirke. Prilikom pritiska dirke, njen zadnji kraj podiže stubić, tako da pero, u kretanju naviše zakači žicu i proizvede zvuk. Istovremeno, tim dodirom sa žicom izbacuje se jezičak iz osnovnog položaja kako pero ne bi pri silaznom kretanju stubića, ponovo zakačilo žicu i time proizvelo novi zvuk. Pri otpuštanju dirke, stubić koji je na donjem kraju opterećen olovom, vraća se u prvobitni položaj, pri čemu posebna opruga vraća i jezičak u vertikalan stav. Pošto se pri vrhu stubića nalazi i mali prigušivač, povratkom u prvobitni položaj on dotiče žicu i prekida njeno treperenje.

Važna karakteristika čembala je u tome da se snagom udara u dirku ne može uticati na jačinu proizvedenog tona, jer trzaj koji čini pero vrlo malo zavisi od te snage. Zato se dinamičke promene mogu ostvariti jedino uključivanjem raznih registara. Pojam registra u ovom smislu ne odnosi se na određeni instrumentalni tonski opseg već predstavlja na određeni način ostvaren zvuk posebnih odlika u dinamici ili tonskoj boji [10]. Ukoliko instrument ima dva manuala, na gornjem manualu svakoj dirci odgovaraju po dve žice, gde je jedna štimovana na nominalnu tonsku visinu dirke, a druga za oktavu više. Na donjem manualu svaka dirka takođe ima po dve odgovarajuće žice, jednu na nominalnoj tonskoj visini, a drugu za oktavu niže štimovanu. Time, za svaki ton postoje u instrumentu po četiri žice. Osim registara za razne dinamičke promene, postoje i takvi koji zvuku instrumenta daju novu boju, sličnu tonu laute ili harfe.

Još jedna od razlika između čembala i klavira jeste značajno bogatije prisustvo harmonika u tonovima čembala u odnosu na klavir. Na Sl. 1. prikazani su usrednjeni spektri tonova klavira i čembala u okviru celog frekvencijskog opsega. Sa slike se može uočiti da zbog prisustva većeg broja harmonika usrednjeni spektar čembala je izdignutiji u odnosu na spektar klavira za otprilike 30 dB na višim frekvencijama.



Sl. 1. Usrednjeni spektri klavira i čembala u okviru celog frekvencijskog opsega.

III. METODOLOGIJA PROCENE SPEKTRA I ODREĐIVANJA KOEFICIJENTA INHARMONIČNOSTI

Kako bi se pojam inharmoničnosti tona mogao objektivizirati, stvorena je formulacija na osnovu koje se može odrediti frekvencija k-tog harmonika odsviranog tona kao:

$$f_k = kf_0\sqrt{1+k^2B}, \tag{1}$$

gde je k redni broj harmonika, f_0 osnovna frekvencija tona, a B koeficijent inharmoničnosti nastao pri sviranju tona [3]. Koeficijent inharmoničnosti B se može izračunati na osnovu (1), gde bi se na osnovu spektra realnog signala odredilo prvih k harmonika odsviranog tona, a zatim na osnovu njih procenila vrednost B. Jedan od predloženih algoritama za automatsku procenu koeficijenta inharmoničnosti je PFD algoritam [8]. Iterativni deo PFD algortima se zasniva na izračunavanju devijacije frekvencije parcijala, gde se proračunava razlika niza frekvencija parcijala dobijenih iz procene spektra i odgovarajućeg niza frekvencija parcijala dobijenih na osnovu jednačine (1) za vrednost B u tekućoj iteraciji. Na osnovu trenda devijacije frekvencije, donosi se odluka da li se koeficijent B povećava ili smanjuje [8]. Sam algoritam se pokazao kao dosta robustan u situacijama kada u spektru signala nedostaje neki harmonik ili kada je frekvencija pojedinog parcijala pogrešno procenjena. Međutim, u prethodnom istraživačkom radu ispostavilo se da za pojedine tonove iz registra klavira dolazi do loše procene koeficijenta *B* [9]. Razlog za lošu procenu koeficijenta *B* leži u početnom koraku samog algoritma gde se vrši procena spektra i određuju vrednosti prvih *k* frekvencija parcijala. Sama procena frekvencija prvih *k* harmonika može se realizovati na različite načine, DFT analizom [8], [12], tehnikama zasnovanim na inharmoničnim češljastim filtrima [5], [13], kao i AR modelima [14]. U ovom radu razmatran je uticaj procene frekvencija harmonika na tačnost procene parametra *B* na primeru snimljenih tonova jednog čembala. Kao metode procene spektra korišćene su metode: DFT analiza i AR modelovanje.

A. Metoda 1 – DFT analiza

DFT analiza predstavlja klasičan pristup u proceni spektra signala. U obradi audio signala, zbog nestacionarnosti signala, često se koristi analiza po segmentima, odnosno „prozorima“. Dodatno, procena tačne pozicije komponenti u spektru se popravlja dodavanjem nula, odnosno računanjem DFT-a u više tačaka nego što je dužina segmenta signala. Dužina prozora, tip prozorske funkcije kao i broj tačaka u kojima se DFT računa su parametri koji utiču na tačnost procene frekvencija harmonika.

Kao ulazni podaci algoritma za procenu koeficijenta *B* koriste se odbirci audio signala snimljenog tona čembala.

Na početku analize, signal je prozorovan *Hamming*-ovom prozorskom funkcijom. Naredni korak je računanje spektra prozorovanog signala pomoću DFT. Kako bi frekvencijska rezolucija spektra signala bila dobra, neophodno je niz dopuniti nulama do vrednosti 2^{16}, a zatim izvršiti izračunavanje spektra signala u 2^{16} tačaka. Dodavanjem nula postignuto je lakše procenjivanje spektralnih komponenti signala.

Za posmatrani audio signal na samom početku vrši se procena osnovne frekvencije tona. Osnovna frekvencija tona se određuje tako što se za osnovnu frekvenciju tona proglasi frekvencija koja odgovara najvećoj amplitudi u spektru za predloženi opseg, definisan na osnovu očekivanih frekvencija za ton koji se analizira. Kod tonova koji pripadaju nižem registru čembala, ova metoda za procenu osnovne frekvencije tona daje loše rezultate. Razlog za lošu procenu leži u tome što u predloženom opsegu za datu osnovnu frekvenciju postoji nekoliko frekvencija čije su pripadajuće amplitude sličnih vrednosti. U tom slučaju, da bi se izbeglo pogrešno procenjivanje osnovne frekvencije za tonove iz nižeg registra procenjuje se frekvencija prvog harmonika tona.

S obzirom na to da je očekivano da se harmonici nalaze na celobrojnim umnoščima osnovne frekvencije tona, moguće je izvršiti manipulaciju nad spektrom u cilju smanjivanja neinformativnih delova spektra signala. U tom kontekstu, spektar se deli na podopsege gde se iz svakog podopsega vrši selekcija 10 spektralnih komponenti sa najvećim amplitudama. Širina podopsega je definisana kao $3f_1$, gde f_1 predstavlja procenjenu frekvenciju prvog parcijala tona. Deo procene spektra, u slučaju primene DFT analize se završava na početku samog iterativnog algoritma, gde se još jednom vrši selekcija spektralnih komponenti. Nad već redukovanihm spektrom signala, vrši se selekcija frekvencija parcijala tona f_k koje odgovaraju maksimalnoj amplitudi spektra unutar zatvorenog intervala. Donja granica intervala predstavlja razliku frekvencije parcijala tona f_k i pomeraja Δf, dok gornja granica intervala predstavlja zbir frekvencije parcijala tona f_k i pomeraja Δf. Za vrednost pomeraja Δf u algoritmu korišćena je vrednost $0.4f_1$.

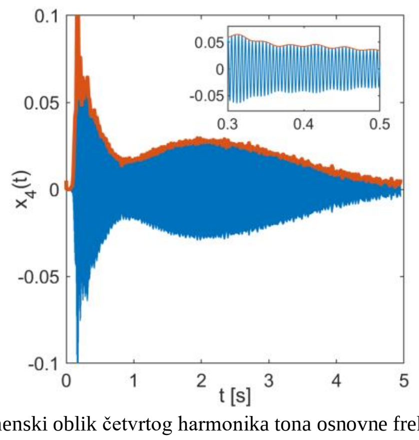
B. Metoda 2 – AR model

AR modelovanje podrazumeva da se na osnovu segmenata signala procene koeficijenti modela. Znajući koeficijente modela, za signale kod kojih su spektralne komponente izražene moguće je izvršiti procenu frekvencija spektralnih komponenti. Iako su za modelovanje audio signala, a posebno tonova muzičkih instrumenata, u literaturi preporučeni ARMA modeli [14], za samu procenu spektra AR modeli su se pokazali kao dovoljno dobar metod. Ujedno, AR modeli su jednostavniji za implementaciju jer se koeficijenti modela mogu izračunati *Yule-Walker*-ovim jednačinama [15]. Podešavanjem reda modela moguće je uticati na tačnost procene spektralnih komponenti.

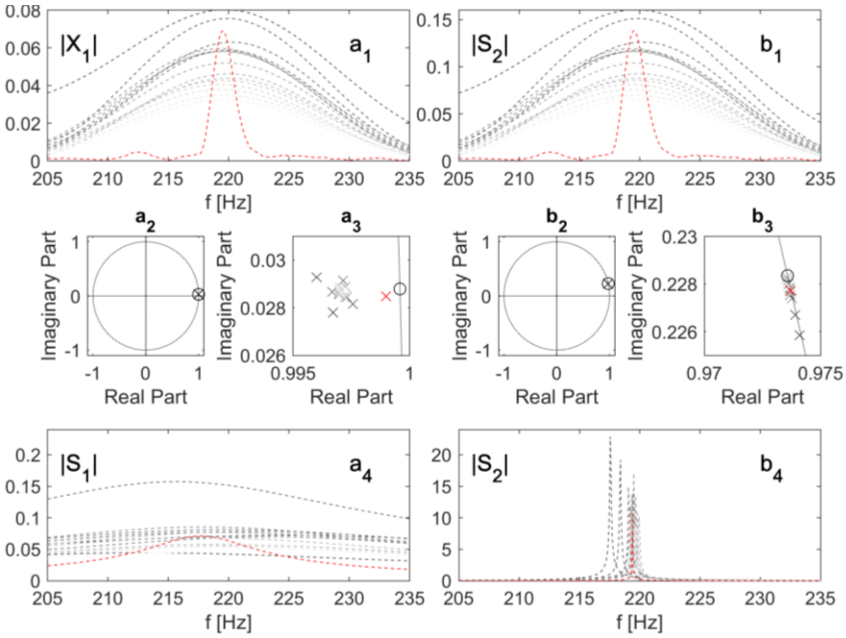
Kao i za slučaj procene spektra metodom DFT analize ulazni podaci algoritma su odbirci audio signala snimljenog tona čembala. Na samom početku potrebno je odrediti red AR modela. Red je empirijski procenjen i variran u zavisnosti od frekvencije pojedinih tonova iz opsega čembala. Zatim je spektar signala podeljen na podopsege čija je širina adekvatna za tri harmonika. Svaki od podopsega je „spušten“ u osnovni opseg i zatim je izvršeno filtriranje i decimacija [14]. Kako faktor decimacije zavisi od osnovne frekvencije tona, korišćene su različite vrednosti faktora decimacije. Krajnji rezultat metode AR modelovanja je procenjivanje frekvencija spektralnih komponenti na osnovu položaja dominantnih polova.

C. Analiza uticaja različitih parametara na procenu spektra

Kako bi se ispitao uticaj različitih parametara na procenu koeficijenta inharmoničnosti biće izvršena analiza signala izdvojenog četvrtog harmonika tona nominalne osnovne frekvencije 55 Hz čiji je segment prikazan na Sl. 2. Signal je izdvojen alatom zasnovanom na digitalnoj komplementarnoj filtarskoj banci, pogodnoj za analizu audio signala [16].



Sl. 2. Vremenski oblik četvrtog harmonika tona osnovne frekvencije 55 Hz.



Sl. 3. Uticaj različitih parametara na procenu frekvencije četvrtog harmonika tona nominalne osnovne frekvencije f_{0N}=55 Hz. a1) DFT na frekvenciji odabiranja f_s a2) Polovi AR modela računati na frekvenciji odabiranja f_s a3) Zumiran detalj slike a2), polovi AR modela, a4) Spektar računat na osnovu modela čiji su polovi prikazani na a2) i a3). b1) DFT na decimiranom signalu. b2) Polovi AR modela računati na decimiranom signalu b3) Zumiran detalj slike b2), polovi AR modela b4) Spektar računat na osnovu modela čiji su polovi prikazani na b2) i b3).

Na Sl. 2 je pored vremenskog oblika četvrtog harmonika crvenom bojom prikazana i njegova anvelopa, dobijena kao moduo Hilbertove transformacije signala. Iako je ovako filtriran signal jednostavniji za analizu, već je na osnovu njegovog vremenskog oblika moguće uočiti određene probleme koji se dodatno usložnjavaju prisustvom većeg broja harmonika. Jedan od problema, koji se može uočiti vizuelnom inspekcijom anvelope harmonika jeste da izdvojeni harmonik nije idealna sinusoida. Činjenica da različiti harmonici istog tona imaju različite anvelope dovodi do toga da u većoj ili manjoj meri njihove anvelope utiču na tačnost procene frekvencije harmonika, a samim tim i na tačnost procene koeficijenta inharmoničnosti. Takođe, signal sa Sl. 2 je u velikoj meri „čist", jer su filtriranjem potisnuti ostali harmonici, kao i eventualni dodatni šumovi. Kada se analizira ceo ton, to nije slučaj, pa analiza postaje složenija. Ukoliko je cilj da se analiza automatizuje, odnosno bude primenljiva za celokupni registar čembala, postupak se dodatno usložnjava.

Na Sl. 3 prikazani su rezultati dobijeni procenom spektra pomoću dve pomenute metode i za različite parametre za slučaj signala sa Sl. 2. Rezultati sa Sl. 3 koji su označeni slovnim indeksom a) dobijeni su na osnovu obrade signala sa zadatom frekvencijom odabiranja od 48 kHz koja odgovara frekvenciji odabiranja kojom su snimani tonovi čembala. Dok se rezultati koji su označeni slovnim indeksom b) odnose na rezultate dobijene nakon decimacije sa faktorom 8.

U prvom redu Sl. 3, odnosno na graficima a1) i b1) prikazani su amplitudski spektri, rezultati dobijeni DFT analizom. Crnom i sivom bojom je naznačena spektralna komponenta računata na segmentu trajanja 0.1 s sa preklapanjem od 75% i dopunjavanjem nulama do rezolucije od 0.1 Hz. Dok je crvenom bojom rezultat dobijen na osnovu DFT-a sračunatog za prvu sekundu snimljenog signala. Korišćena je *Hamming*-ova prozorska funkcija. Prikazani su rezultati za nekoliko segmenta signala. Može se uočiti da se pozicije maksimuma, odnosno procene frekvencije harmonika, menjaju od prozora do prozora. Poređenjem Sl. 3 a1) i b1) može se uočiti da decimacija nema značajan uticaj na estimaciju frekvencije.

U drugom redu Sl. 3, na graficima a2), a3), b2) i b3), prikazani su polovi AR modela. Polovi su računati tako da modeluju signal četvrtog harmonika nakon Hilbertove transformacije, pa su dobijeni modeli sa kompleksnim koeficijentima. S obzirom da se analizira samo jedan harmonik, red modela je 1. Na grafiku a2) prikazani su polovi AR modela računati sa frekvencijom odabiranja f_s, dok su na grafiku b2) prikazani polovi AR modela računati na decimiranom signalu. Kružićem je prikazana pozicija na krugu koja odgovara nominalnoj frekvenciji četvrtog harmonika od 220 Hz. Frekvencija od 220 Hz ne predstavlja „tačnu" vrednost frekvencije četvrtog harmonika, jer tačna frekvencija odsviranog tona nije unapred poznata, već je poznata njena vrednost do nivoa raspona tona. S obzirom na izraz (1) bilo bi očekivano da je dobijena frekvencija zapravo nešto viša od 220 Hz, ali na osnovu prikazanog rezultata može se zaključiti da je osnovna frekvencija odsviranog tona nešto niža od nominalne vrednosti od 55 Hz. Na a3) i b3) prikazani su zumirani detalji sa grafika a2) i b2), respektivno. Crni i sivi krstići predstavljaju polove računate po prozorima dužine 0.1 s, gde se na osnovu slike može uočiti da se pozicija pola menja od prozora do prozora. Crvenom bojom je prikazan pol modela koji je dobijen na osnovu 1 s signala. Ukoliko se izvrši poređenje grafika a3) i b3) može se uočiti da u ovom slučaju decimacija doprinosi boljoj proceni utoliko što je pol pozicioniran bliže jediničnom krugu, pa je njegov efekat izraženiji. Intuitivno je jasno da decimacija doprinosi boljim rezultatima jer, se jedne strane, odbačen deo spektra koji nije od interesa, a s druge strane je relativan odnos frekvencije od interesa i frekvencije odabiranja manji, što se vidi i kroz razliku položaja nominalne frekvencije na jediničnom krugu, grafici a2) i b2). U poslednjem redu na graficima a4) i b4), prikazani su spektri dobijeni na osnovu AR modela koji verifikuju zaključak o uticaju decimacije na bolju procenu frekvencije [14].

IV. EKSPERIMENTALNI REZULTATI I DISKUSIJA

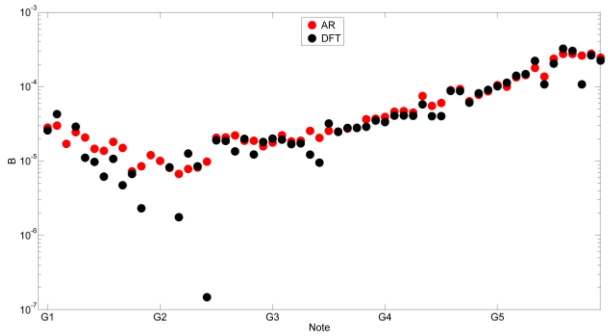
Na osnovu analize izložene u prethodnom delu, za dalju diskusiju usvojene su dve metode za procenu spektra, dok je za samu procenu koeficijenta inharmoničnosti korišćen iterativni PFD algoritam [10]. Korišćene su metode DFT analiza i AR modelovanje. Obe metode za procenu spektra su definisane tako da imaju odliku automatske obrade, odnosno da daju vrednosti koeficijenta inhrmoničnosti za celokupni registar čembala. Algoritam je testiran na bazi formiranoj za potrebe ovog rada. U bazi se nalaze tonovi snimljeni na 5 različitih registara jednog čembala.

A. Baza snimaka

Za potrebe ovog rada napravljena je baza koju čini 310 snimaka tonova odsviranih u 5 različitih registra čembala. Tonovi čembala su snimljeni u prostorijama Fakulteta muzičke umetnosti u Beogradu. U analizi uticaja tačnosti estimacije frekvencija harmonika na procenu koeficijenta inharmoničnosti korišćeni su tonovi iz prva dva registra. Izabrani su tonovi iz ova dva registra, jer njihova upotreba služi za pojačanje izvođačke dinamike, a ne za menjanje boje instrumenta.

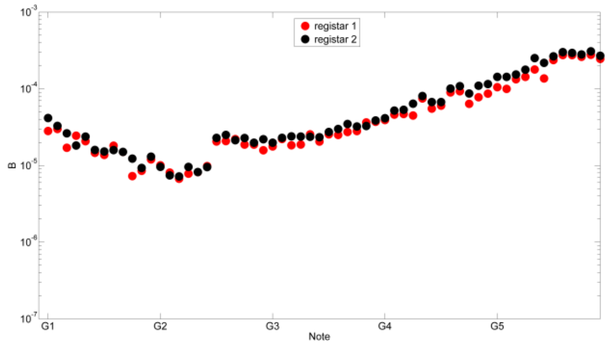
B. Analiza uticaja različitih parametara na tačnost procene koeficijenta inharmoničnosti

Uticaj tačnosti različitih metoda za estimaciju spektra na procenu koeficijenta inharmoničnosti najbolje se može uočiti poređenjem trenda krive koeficijenta *B*. Na Sl. 4 prikazan je grafik koeficijenta inharmoničnosti *B* za celokupan registar čembala procenjen na osnovu dve metode za procenu spektra i to DFT-om i AR modelovanjem. Za tonove iz srednjeg i višeg registra čembala koeficijenti *B* procenjeni na osnovu dve metode se ne razlikuju značajno, dok za slučaj tonova iz nižeg registra dolazi do razilaženja u vrednostima. Razlog za lošiju procenu koeficijenta *B* na osnovu metode DFT jeste u proceni osnovne frekvencije tona. Sam iterativni algoritam iako robustan u slučajevima kada u spektru signala nedostaje neki harmonik ili kada je frekvencija pojedinog parcijala pogrešno procenjena, usled pogrešno procenjene osnovne frekvencije tona kao krajnji rezultat daje loše procenjnu vrednost koeficijenta *B*. Do pogrešno procenjene osnovne frekvencije tona DFT metodom dolazi zbog samog spektra signala, gde je na nižim frekvencijama često osnovni harmonik potisnut u odnosu na ostale harmonike tog tona. Takođe, pri proceni spektra DFT analizom rezolucija po frekvenijskoj osi je ista za ceo opseg koji se analizira, dok se frekvencije osnovnih tonova menjaju logaritamski, te tačnost procene u muzičkom smislu nije ista u celom opsegu. Metoda AR modelovanjem dala je tačnije rezultate za koeficijent *B* . Osnovna frekvencija tona je procenjivana metodom najmanjih kvadrata (*Least square*) nad nekoliko parcijala tona počev od drugog. Time je premošćen problem korišćenje osnovnog parcijala. U najvećem broju snimljenih tonova čembala harmonici nižeg reda su dovoljno jasni, odnosno procena njihovih frekvencija je relativno tačna. Broj harmonika za procenu osnovne frekvencije, kao i broj harmonika na osnovu kojih je izvršena procena parametra *B* variran je u zavisnosti od visine tona, jer je za tonove nižih frekvencija broj jasno prisutnih harmonika u snimljenom tonu veći.



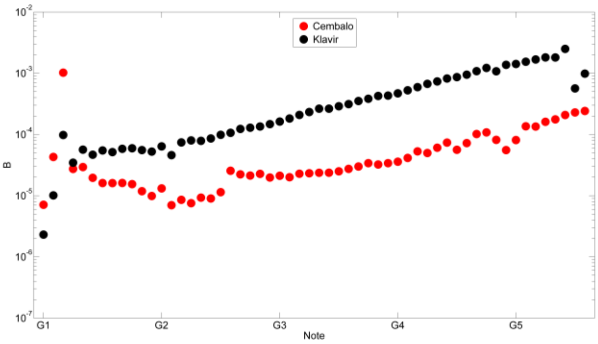
Sl. 4. Koeficijent inharmoničnosti čembala dobijen metodama: AR i DFT.

Na Sl. 5 prikazana je kriva koeficijenta inharmoničnosti *B* za tonove iz prvog i drugog registra čembala. Koeficijenti *B* su procenjeni na osnovu estimiranih frekvencija harmonika pomoću AR modela. Trend krive koji je očuvan i za tonove iz prvog i drugog registra čembala verifikuje tačnost estimacije frekvencija harmonika koja je dobijena na osnovu AR modela.



Sl. 5. Koeficijenti inharmoničnosti čembala iz prvog i drugog registra.

S obzirom da se tonovi čembala u spektru odlikuju većim brojem harmonika u odnosu na tonove klavira istih frekvencija, izvršeno je poređenje koeficijenta *B* za tonove čembala u odnosu na koeficijent *B* za tonove klavira posmatrane na istom frekvencijskom opsegu. Na Sl. 6 prikazani su koeficijenti inharmoničnosti tonova čembala iz prvog registra i polukoncertnog klavira Steinway D7. Na osnovu slike se može zaključiti da postoji sličan trend krive koeficijenata *B* za oba instrumenta, iako se ova dva instrumenta razlikuju na osnovu načina nastanka tona. Kod tonova čembala, vrednosti koeficijenta *B* su manje za red veličine. Pored velikog broja harmonika tonova čembala u odnosu na klavir, postoji još jedna razlika između ova dva instrumenta koja doprinosi velikim odstupanjima koeficijenta inharmoničnosti. Kod klavira u većinskom frekvencijskom opsegu jednom tonu odgovaraju tri žice, dok čembalo za svaki ton ima jednu. S obzirom na postojanje tri žice kod klavira, koje se vide u spektru signala, pri proceni spektra uvek se donosi odluka na osnovu jedne spektralne komponente na osnovu čije frekvencije se ulazi u dalju procenu koeficijenta inharmoničnosti. Tim fenomenom se otvaraju nove potrebe u smislu modelovanja uticaja tri žice na procenu spektra, a samim tim i procenu koeficijenta inharmoničnosti.



Sl. 6. Koeficijent inharmoničnosti *B* tonova čembala i klavira.

V. ZAKLJUČAK

U ovom radu razmatran je uticaj procene frekvencija harmonika na tačnost procene koeficijenta inharmoničnosti. Formirana je baza snimaka koja se sastoji od preko 300 tonova odsviranih na čembalu. Čembalo je izabrano kao instrument za analizu, jer se spektri njegovih tonova odlikuju velikim brojem harmonika. Prisutnost velikog broja parcijala u spektru tona je značajna, zbog činjenice da su fizičke pojave koje opisuju fenomen inharmoničnosti uočljivije na višim parcijalima tona. Sem toga sa prisustvom većeg broja harmonika iterativni algoritam bolje procenjuje koeficijent inharmoničnosti. Izvršena je uporedna analiza dve metode za procenu spektra signala i to procena spektra na osnovu AR modela i metodom DFT. Pokazano je da je procenu frekvencije harmonika bolje vršiti na osnovu AR modela, nego metodom DFT, jer trend krive procenjenih koeficijenata inharmoničnosti u tom slučaju ima konzistentnije ponašanje. Poređenjem, koeficijenta inharmoničnosti tonova klavira i čembala uočeno je da na njihove vrednosti između ostalog utiče i broj žica koji je potreban za sviranje jednog tona. Dalja unapređenja postupka za procenu koeficijenta inharmoničnosti biće predlog rešenja za modelovanje uticaja prisustva više žica na koeficijent inharmoničnosti. Izvedeni zaključci su korisni za buduća istraživanja gde će se na ovaj način procenjeni koeficijenti inharmoničnosti koristiti u sintezi i modelovanju tonova žičanih instrumenata koje će predstavljati osnovu subjektivnih testova. Cilj subjektivnih testova bi bio upotpunjavanje istraživanja o uticaju inharmoničnosti žičanih instrumenata na ljudsku percepciju tonova.

ZAHVALNICA

Ovaj rad je realizovan u okviru projekta TR36026 koji finansira Ministarstvo prosvete, nauke i tehnološkog razvoja Republike Srbije. Zahvaljujemo se kolegama sa Fakulteta muzičke umetnosti na ustupljenim muzičkim instrumentima i formiranju baze snimaka.

LITERATURA

[1] J. W. Strutt, *The Theory of Sound*, Macmillan 1894.
[2] R.W. Young, "Inharmonicity of Plain Wire Piano Strings", *J. Acoust. Soc. Am.*, vol. 24, no. 3, pp. 267–273, 1952.
[3] H. Fletcher, E. D. Blackham, and R. Stratton, "Quality of piano tones", *J. Acoust. Soc. Am.*, vol. 34, no. 6, pp. 749–761, 1962.
[4] H. Fletcher, "Normal Vibration Frequencies of a Stiff Piano String", *J. Acoust. Soc. Am.*, vol. 36, no. 1, pp. 203–209, 1964.
[5] A.S. Galembo, and A. Askenfelt, A. "Signal representation and estimation of spectral parameters by inharmonic comb filters with application to the piano," *IEEE Trans. Speech Audio Process*, vol. 7, no. 2, pp. 197–203, 1999.
[6] A. Askenfelt and A.S. Galembo, "Study of the spectral inharmonicity of musical sound," *Acoust. Phys*. vol. 46, no. 2, pp. 121–132, 2000.
[7] A. Klapuri, "Multiple fundamental frequency estimation based on harmonicity and spectral smoothness," *IEEE Trans. Speech Audio Process*, vol. 11, no. 6, pp. 184–194, 2003.
[8] J. Rauhala, H.M. Lehtonen, V. Välimäki, "Fast automatic inharmonicity estimation algorithm" *J. Acoust. Soc. Am.*, vol. 121, EL184, doi: 10.1121/1.2719043, 2007.
[9] T. Miljković, M. Bjelić, D. Šumarac Pavlović, J. Ćertić, "Analiza algoritma za procenu koeficijenta inharmoničnosti različitih klavira", Zbornik radova 64. Konferencije ETRAN, pp. AK1.2.1 - AK1.2.6, Beograd, Sep, 2020.
[10] D. Despić, *Muzički instrumenti*, Univerzitet umetnosti Beograd, 1979.
[11] A. Bond, *A Guide to the Harpsichord*, Portland, USA, Amadeus Press, 1997.
[12] S. Dixonb, M. Mauch, and Dan Tidhar "Estimation of harpsichord inharmonicity and temperament from musical recordings", *J. Acoust. Soc. Am.*, vol. 131, no. 1, Pt. 2, pp. 878–887, Jan 2012.
[13] H.M. Lehtonen "Analysis of Piano Tones using an Inharmonic Inverse Comb Filter", Proc. of the 11th Int. Conference on Digital Audio Effects (DAFx-08), Espoo, Finland, September 1-4, 2008.
[14] P.A.A. Esquef, M. Karjalainen, and V. Välimäki, "Frequency-Zooming ARMA Modeling for Analysis of Noisy String Instrument Tones", *EURASIP J. Adv. Signal Process*, pp. 953-967, 2003.
[15] S. Haykin, *Adaptive Filter Theory*, 4th ed. New Jersey, USA, Prentice Hall, 2002.
[16] J. Ćertić, D. Šumarac Pavlović, I. Salom, "Softverski paket za obradu i analizu audio signala", *TELFOR 2010*, pp. 1269-1272, Beograd, Nov., 2010.

ABSTRACT

Inharmonicity is a phenomenon that occurs in musical instruments that are theoretically considered harmonic, and it is the deviation of the frequencies of the tone partials from the integer multiple of the fundamental frequency. For stringed musical instruments, the inharmonicity coefficient is defined in the literature as a measure of deviation from the ideal case. In the previous research work, it was shown that an automatic algorithm for estimating the inharmonicity coefficient of piano tones, for individual tones from the register doesn't perform a good estimation of the inharmonicity coefficient. In this paper, the accuracy impact of harmonics frequency estimation on the estimation of inharmonicity coefficient is considered. Comparative analysis of two methods for signal spectrum estimation was performed, namely spectrum estimation based on AR model and DFT method. Testing of the proposed methods for spectrum estimation was performed on real harpsichord tones. It was found that based on the trend of the inharmonicity coefficient calculated using both spectrum estimation methods, the AR model method is superior and performs a more accurate estimation of the inharmonicity coefficient over the whole range of tones of interest.

Impact of harmonic frequency estimation on the estimation of harpsichord inharmonicity coefficient

Tatjana Miljković, Jovana Damnjanović, Jelena Ćertić, Dragana Šumarac Pavlović

Vreme reverberacije energetskog transformatora

Miloš Bjelić, Bogdan Brković, Mileta Žarković, Tatjana Miljković

Apstrakt— **Energetski transformator je uz generator jedan od najvažnijih elemenata elektroenergetskih sistema. Pouzdanost rada energetskog transformatora direktno je povezana sa njegovim stanjem. Analiza rezultata ispitivanja svodi se na vizuelno poređenje frekvencijskih odziva periodičnih merenja transformatora, dobijenih pomoću specijalizovanih uređaja. Nedostatak ove metode je oslanjanje na inženjersko iskustvo. Objektivniji način detekcije kvarova je izračunavanje matematičkih indikatora zasnovanih na analizi frekvencijskog odziva. Ovi parametri često nisu u stanju da nedvosmisleno ukažu na stepen kvara. Ovaj rad predstavlja novi pristup za otkrivanje unutrašnjih kratkih spojeva transformatora. Predlaže se novi indikator zasnovan na vremenu reverberacije, koje se uobično koristi u akustici. Predloženi parametar zasnovan je na proračunu u vremenskom domenu i omogućava pouzdanu detekciju kratkih spojeva transformatora i tačnu procenu njihovog stepena.**

Ključne reči — **Detekcija kvara, energetski transformator, indikator, obrada signala, vreme reverberacije.**

I. Uvod

Energetski transformator (u daljem tekstu ET) je jedan od najvažnijih i najskupljih elemenata u svakom delu i na svakom nivou napona elektroenergetskog sistema. Pored toga, ET je izuzetno složena komponenta na čiji rad utiču termički, mehanički, električni i hemijski procesi. Neki od ključnih delova transformatora koji su od velike važnosti za njegov pouzdan rad su: magnetni krug, ulje, namotaji (stanje geometrije i stanje električnih veza) i čvrsta izolacija namotaja. Deformacije namotaja obično se otkrivaju metodom analize frekvencijskog odziva (FRA) i *sweep* analizom frekvencijskog odziva (SFRA). Standardna SFRA metoda uključuje poređenje frekvencijskih odgovora uzastopnih periodično sprovedenih testova. Ako se dva uzastopna frekvencijska odziva ne razlikuju, tada se uzdužni parametri R, L i C nisu značajno promenili, što ukazuje da u međuvremenu nije došlo do deformacija namotaja.

U radovima i standardima [1-14] predstavljene su FRA i SFRA metoda i njihov značaj u praćenju i dijagnostici stanja ET. Referenca [1] predstavlja teorijske osnove FRA metode. Razvoj akustičkih metoda pomoću vibracija predstavljen je u [2], gde je pokušano da se rezultati FRA metode dopune i povežu sa postojećim IEEE standardima [3], IEC standardom [4] i iskustvom radne grupe CIGRE [5, 6]. Dobra interpretacija rezultata FRA metode predstavljena je u [7], gde je pokušano da se odrede pojedinačni indikatori i njihove približne granične vrednosti. Metoda koja grupiše rezultate simulacije FRA metode i povezuje ih sa određenom vrstom kvara predstavljena je u [8]. Eksperimenti na ET u praznom hodu i uslovima kratkog spoja, uz primenu FRA metode predstavljeni su u [9, 10]. Detekcija radijalnih deformacija namota pomoću impulsnog napona groma prisutna je u [11]. Uobičajeni statistički indikatori kvarova dati su u [12-14]. U članku [15] je predstavljena nova metodologija za dijagnostiku kvara ET, zasnovana na rezultatima SFRA. Ova metoda upoređena je sa metodom konačnih elemenata pri otkrivanju unutrašnjih i spoljnih kvarova kratkog spoja [16].

Tumačenje rezultata SFRA obično se vrši vizualizacijom slika generisanih od strane specijalizovanih uređaja za takva ispitivanja. Razvoj novih metoda motivisan je potrebom za postizanjem objektivne interpretacije rezultata. Ovaj rad prikazuje neke najčešće korišćene identifikatore koji se primenjuju na dobijene SFRA testove. Takvi identifikatori trebali bi jasnije da kvantifikuju rezultata SFRA metode i omogućiti lakše i nedvosmisleno otkrivanje oštećenja namotaja ET i kvarova magnetnog kruga.

U ovom radu predložena je detekcija kratkih spojeva i određivanje njihove težine vrši se na osnovu vremena reverberacije [17]. Ovaj parametar se najčešće koristi za analizu akustike prostorija. Glavna hipoteza istraživanja u ovom radu je da se broj kratko spojenih namotaja može odrediti na osnovu vrednosti vremena reverberacije. Realizovani eksperimenti obuhvatili su testiranje sa kratkim spojem različitih delova namotaja jednog ET. Merenja su sprovoda pomoću audio kartice. Snimljeni signali se koriste za izračunavanje statističkih parametara koji se obično koriste u literaturi za detekciju kvarova ET. Izvršeno je poređenje između vremena reverberacije i postojećih parametara u pogledu osetljivosti na dužinu kratkih spojeva.

Rad je organizovan na sledeći način. U odeljku 2 dat je pregled najznačajnijih metoda i statističkih parametara koji se koriste za ET dijagnostiku. U ovom poglavlju objašnjen je postupak izračunavanja vremena reverberacije. Eksperimentalna postavka opisana je u poglavlju 3. Rezultati dobijeni korišćenjem zvučne kartice prikazani su u poglavlju 4. Zaključci istraživanja dati su u poglavlju 5.

Miloš Bjelić – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: bjelic@etf.rs).

Bogdan Brković – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: brkovic@etf.rs).

Mileta Žarković – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: mileta@etf.rs).

Tatjana Miljković – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: tm@etf.rs).

II. METODOLOGIJA

A. *FRA i SFRA analiza*

U svrhu otkrivanja mehaničkih kvarova ET, mogu se koristiti različiti ulazni signali:

- odskočni signal (FRA analiza),
- *sweep* signal (SFRA analiza).

Kada se ET terminal pobudi sa nekim od ulaznih signala, meri se odziv na izlazu ET. U ovim merenjima, namotaj ET može se posmatrati kao dvopolna mreža koja sadrži otpornike, kondenzatore i kalemove, gde je otpor obično zanemarljiv. Ako se dva upoređena frekvencijska odziva (periodična testiranja transformatora) ne razlikuju, to znači da se parametri nisu značajno promenili, tj. nije bilo deformacija namotaja od prethodnog merenja. U slučaju mehaničkih oštećenja namotaja, vrednosti ekvivalentnih kondenzatora i kalemova se menjaju, pa se odzivi dva periodična ispitivanja se razlikuju, što ukazuje na kvar. Frekvencijski odziv $H(j\omega)$ se formira kao odnos Furijeove transformacije izlaznog signala V_{output} i Furijeove transformacije ulaznog signala V_{input}:

$$H(j\omega) = \frac{V_{output}(j\omega)}{V_{input}(j\omega)}. \tag{1}$$

Za analizu se najčešće koristi samo amplitudska karakteristika frekvencijskog odziva. Ordinata predstavlja amplitudu u decibelima, a apscisa predstavlja frekvenciju na logaritamskoj skali. Ovaj postupak prelaska u frekvencijski domen u velikoj meri olakšava analizu i pruža nezavisnost od talasnog oblika pobude. Međutim, to povećava računsku i vremensku kompleksnost samog potapka. Poređenjem izračunatih frekvencijskih odziva u uzastopnim ispitivanjima mogu se otkriti različite promene u geometriji namotaja i promene u magnetnom kolu ET. Tumačenje rezultata SFRA oslanja se na ličnu stručnost i može dovesti do različitih zaključaka. Štaviše, konvencionalni SFRA *fingerprint* vrlo je neprecizan u otkrivanju početnih i niskih nivoa mehaničkih kvarova [13]. Da bi se rešili ovi problemi, predloženi su različiti statistički indikatori [6, 12, 13, 14]. Neki od ovih indikatora definisani su u tabeli I. Analiza osetljivosti indikatora može se izvršiti izračunavanjem njihovih vrednosti za različite intenzitete kvara.

TABELA I
KORIŠĆENI STATISTIČKI INDIKATORI IZ LITERATURE

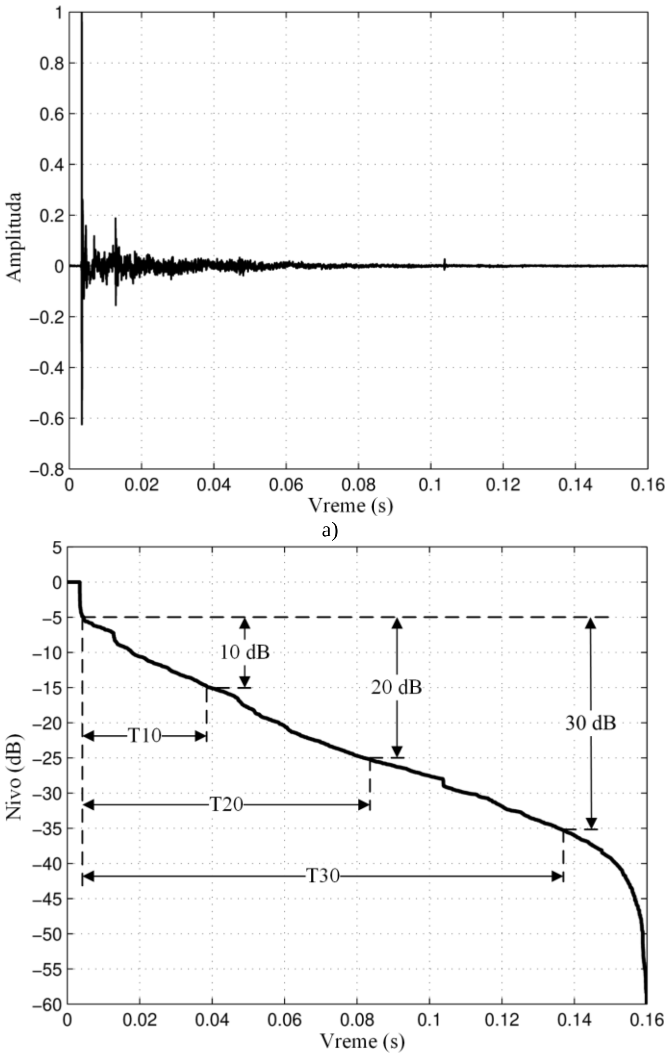
Indikator	Opis
Jaccard Distance (JD)	$\dfrac{\sum_{i=1}^{N}(x_i - y_i)^2}{\sqrt{\sum_{i=1}^{N}(x_i-\bar{x})^2 \sum_{i=1}^{N}(y_i-\bar{y})^2}}$
Absolute sum of logarithmic error (ASLE)	$\dfrac{\sum_{i=1}^{N}\left\|20\log_{10}(y_i) - 20\log_{10}(x_i)\right\|}{N}$
Absolute average difference (DABS)	$\dfrac{\sum_{i=1}^{N}\left\|y_i - x_i\right\|}{N}$

(N – ukupan broj podataka; x_i i y_i – ulazni, odnosno izlazni podaci; $\bar{x}$ i $\bar{y}$ – srednje vrednosti)

B. *Vreme reverberacije*

Linearni vremenski invarijantni sistem (LTI) može se jednoznačno opisati frekvencijskim odzivom ili impulsnim odzivom u vremenskom domenu [18]. Frekvencijski odziv sistema se dobija primenom Furijeove transformacije na impulsni odziv u vremenskom domenu. Impulsni odziv se može dobiti pobuđivanjem sistema Dirakovim impulsom. Međutim, generisanje Dirakovog impulsa u praksi nije moguće, pa se umesto toga moguće koriste jedinični impuls. Češće se koriste indirektne metode kao što su *sweep* signal (*chirp*) i MLS sekvenca [19]. Impulsni odziv se obično koristi u oblastima obrade signala, automatskog upravljanja i akustike.

Impulsni odziv prostorije meri se pomoću zvučnika i mikrofona. Soba se može posmatrati kao LTI sistem, gde signal mikrofona predstavlja izlaz sistema, a signal koji se emituje iz zvučnika predstavlja ulaz sistema. Impulsni odziv prostorije predstavlja njenu „ličnu kartu“, jer omogućava određivanje gotovo svih akustičkih parametara prostorije. Primer impulsnog odziva prostorije prikazan je na slici 1 a.



Sl. 1. Određivanje vremena reverberacije za prostoriju: a) Impulsni odziv, b) Kriva opadanja i princip određivanja T10, T20 i T30

Struktura impulsnog odziva je vrlo složena i njegov talasni oblik ne pruža mnogo informacija o prostoriji. Zvučno polje obično se može analizirati jednostavnim posmatranjem globalnog oblika impulsnog odziva. Zbog toga se impulsni odziv najčešće karakteriše tzv. krivom opadanja zvuka (reverberaciona kriva) [17]. Da bi se smanjio uticaj fluktuacija krive, koristi se tzv. integrisani impulsni odziv, poznat i kao *Schroeder*-ova kriva [17]. Ovaj pristup zasnovan je na uvođenju krive opadanja $L_R(t)$, definisane kao:

$$L_R(t) = \int_t^{\infty} h^2(t)\,dt = \int_0^{\infty} h^2(t)\,dt - \int_0^{t} h^2(t)\,dt, \qquad (2)$$

gde je $h(t)$ impulsni odziv. Na slici 1 b) prikazana je kriva opadanja impulsnog odziva, prikazanog na slici 1 a.

Najčešće korišćeni parametar u akustici prostoriji je vreme reverberacija (u oznaci T60). Vreme reverberacije je definisano kao vreme potrebno da se kriva reverberacije smanji za 60 dB od početne vrednosti, što odgovara smanjenju od milion puta na linearnoj skali. Može se smatrati da nakon vremena reverberacije u sobi nema zvučne energije. Dinamiku od 60 dB je teško postići u praksi, zbog čega se T60 određuje indirektno pomoću parametara T10, T20 ili T30. Tumačenje ovih parametara ilustrovano je na slici 1 b [17]. Vreme reverberacije T60 izračunava se množenjem T30 sa dva, T20 sa tri ili T10 sa šest.

U ovom radu se ispituje mogućnost korišćenja vremena reverberacije kao indikatora za određivanje stanja ET. Za razliku od akustike prostorije, gde su pobuda i impuls zvučni signali (zvučni pritisak), pobuda i impulsni odziv ET-a su električni signali (napon odnosno struja). Iako su izmerene veličine različite prirode, princip izračunavanja vremena odjeka važi u oba fizička domena. U opštijem smislu, vreme reverberacije može se tumačiti kao vreme smirivanja impulsnog odziva.

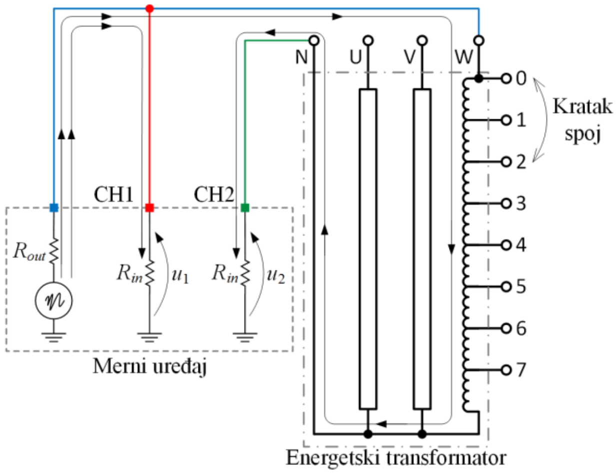
III. Postavka eksperimenata

Eksperimentalna postavka je šematski prikazana na slici 2. ET koji se koristi u postavci je proizveden u Laboratoriji za visoki napon Elektrotehničkog fakulteta, Univerziteta u Beogradu. Parametri transformatora dati su u tabeli II.

TABELA II
PARAMETRI KORIŠĆENOG ENERGETSKOG TRANSFORMATORA

Parametar	Vrednost
Nazivni napon	10/0.4 kV/kV
Nazivna struja	23/557 A/A
Frekvencija	50 Hz
Tip veze	Ynyn0
Broj faza	3

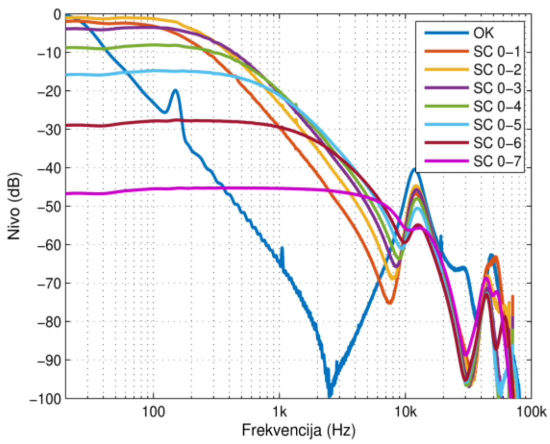
Ovaj ET ima sedam dvostrukih namotaja za sa pristupačnim kontaktima na namotaju faze W, označenim sa 0 do 7 na slici 2. Takva konstrukcija omogućava kratki spoj različitih delova namotaja. U zavisnosti od veza između ovih kontakta, stepen kvara može se menjati, pri čemu kratki spojevi između susednih kontakata odgovaraju najmanjem, a kratki spoj između kontakata 0 i 7 odgovaraju najvećem stepenu kvara.



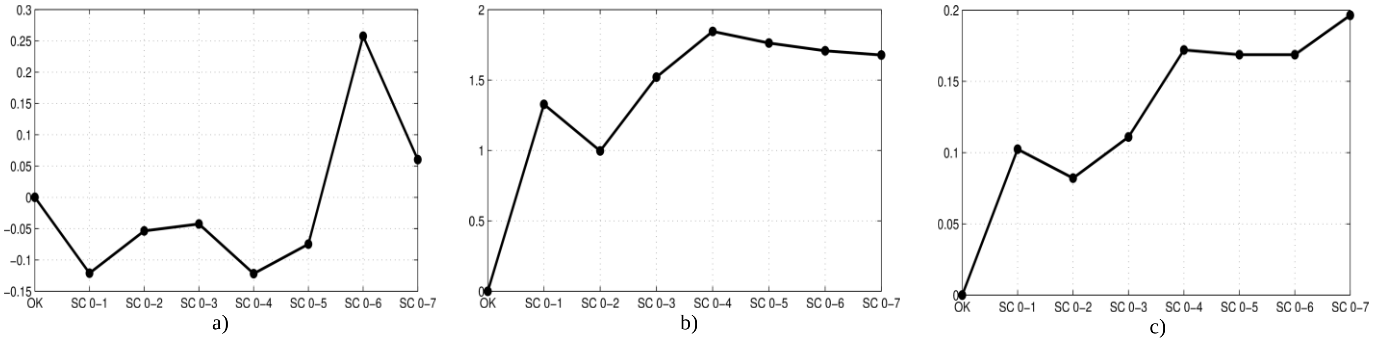
Sl. 2. Šematski prikaz eksperimentalne postavke

Fazni W namotaj se pobuđuje pomoću izvora promenljive frekvencije (plava linija). Ulazni napon (u_1) meri se između fazne W i zemlje (Kanal 1 - crvena linija). Napon koji odgovara odzivu (u_2) se meri između neutralne tačke transformatora i zemlje (Kanal 2 - zelena linija). Na osnovu ova dva merenja može se dobiti frekvencijski odziv namotaja. Putanje struja u kolu su označene crnim linijama sa strelicama koje pokazuju smer struje.

Ukupno je izvršeno osam merenja pomoću svakog, jedno na zdravom ET (označeno sa OK) i sedam na neispravnom ET (označeno kao SC – *Short Circuit*). Kvarovi su nastajali formiranjem kratkog spoja između kontakta 0 i kontakta 1 do 7 (videti sliku 2). Na taj način formirani su kvarovi različite težine. Pobudni signal je bio logaritamski *sweep* signal. Kao merni uređaj korišćena je audio kartica Steinberg UR22 [21]. Ovaj uređaj predstavlja eksternu zvučnu karticu koja se obično koristi za snimanje audio signala. Cena uređaja je oko 150$, što je nekoliko redova veličine niže u poređenju sa profesionalnim SFRA analizatorom. Vreme potrebno za jedno merenje je oko 10 s. Maksimalna frekvencija odabiranja zvučne kartice je 192 kHz. Na osnovu toga ovaj uređaj omogućava merenje frekvencijskog odziva u opsegu od 20 Hz do oko 100 kHz. Merenje je realizovano korišćenjem aplikacije, realizovane u programskom paketu Matlab.



Sl. 3. Frekvencijski odziv dobijen pomoću audio kartice



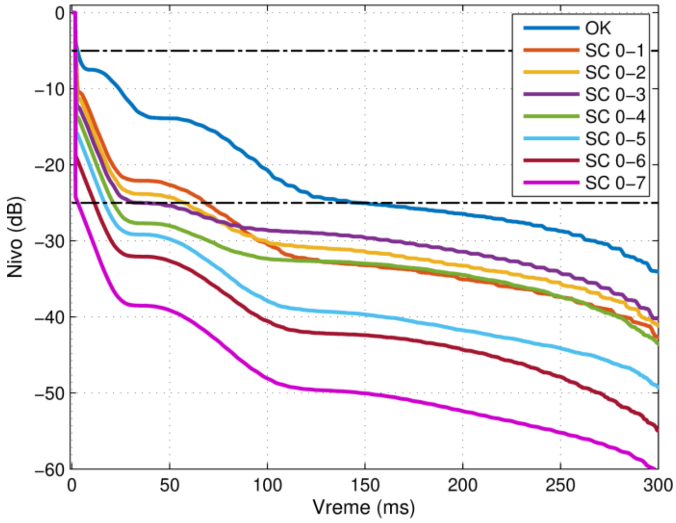
Sl. 4. Statistički indikatori (tabela I): a) Jaccard Distance (JD); b) Absolute sum of logarithmic error (ASLE); c) Absolute average difference (DABS)

IV. Rezultati i diskusija

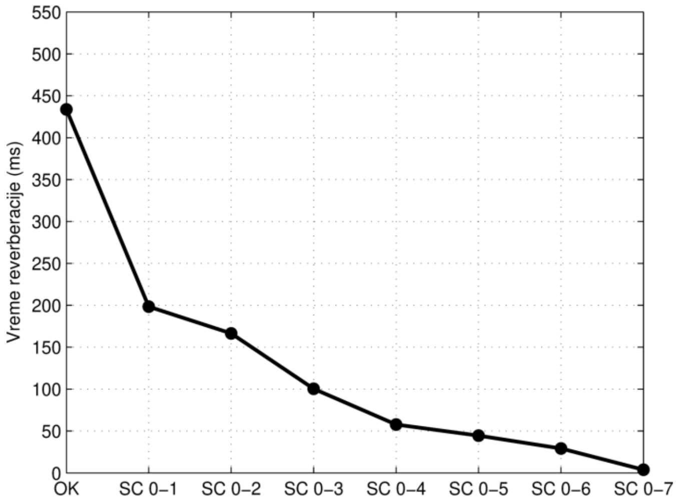
Frekvencijski odziv ET-a sa zdravim namotajem (plava boja) i sa kratkim spojevima (druge boje) prikazani su na slici 3. Povećani nivo u frekvencijskom odzivu primećuje se kada su prisutni kvarovi na transformatoru, što se može pripisati promenjenim vrednostima parametara namotaja (induktivnosti i kapaciteta). ET se može tretirati kao složena *RLC* mreža, sa kratkim spojevima koji utiču na ekvivalentnu impedansu namotaja [22]. Ovaj efekat je najizraženiji na niskim frekvencijama kada su dominantne induktivnosti namotaja. Smanjenje impedanse namotaja dovodi do povećanja napona na CH2 (slika 2), povećavajući tako frekvencijski odziv. Rezonantna frekvencija zadržava vrednost od oko 10 kHz bez obzira na stanje namotaja. Frekvencijski odzivi sa slike 3 dobijeni su korišćenjem izmerenih signala iz kanala 1 (pobuda) i kanala 2 (odziv). Na osnovu ovih signala izračunavaju se statistički pokazatelji definisani u tabeli I. Izračunate vrednosti su prikazane na slici 4 za sve analizirane slučajeve.

Vrednosti parametra JD date su na slici 4 a. Ovaj indikator je jednak 0 za zdrav ET i kreće se od -0.13 do 0.25 kada su prisutni kratki spojevi na namotaju ET. Na osnovu vrednosti ovog indikatora ne može se utvrditi da li na ET postoji kvar. Jasno je da se onda ne može utvrditi ni stepen kvara, odnosno dužina namotaja koja je u kratkom spoju. Indikatori ASLE i DABS jasno ukazuju na kratak spoj namotaja ET. To omogućava vrlo pouzdano otkrivanje kvara. Kao što je vidljivo sa slike 4 b i c, nijedan od dva indikatora se ne menja monotono, pa se stepen kvara ne može pouzdano utvrditi. Ipak, ova dva indikatora imaju očiglednu prednost u poređenju sa JD indikatorom prikazanim na slici 4 a. Oni pružaju najpouzdaniju detekciju kvara i mogu u određenoj meri ukazati na ozbiljnost kvara.

Impulsni odziv ET (tj. namotaja) izračunat je za svaki eksperiment na osnovu snimljenih ulaznih i izlaznih signala. Krive opadanja nivoa signala, koje odgovaraju svim analiziranim scenarijima, dobijaju se na osnovu izraza (3), koristeći izračunate impulsne odzive. Krive opadanja integrisanog impulsnog odziva prikazane su na slici 5. Sve krive pokazuju slično ponašanje tokom prvih 15 ms. Nakon ovog intervala, kriva koja odgovara ispravnom transformatoru sporije opada u odnosu na krive koje odgovaraju ET sa kvarom. Kako se ozbiljnost kvara, tj. broj kratko spojenih navojaka, povećava tako se povećava i nagib krivih opadanja.



Sl. 5. Krive opadanja impulsnih odziva transformatora



Sl. 6. Vreme reverberacije transformatora za različite kvarove

Vreme reverberacije (T60) izračunato je prema slici 2 b. Za izračunavanje T60 korišćen je dinamički opseg od 20 dB, tj. parametar T20. Vrednosti vremena reverberacije dobijene za ispravan ET i analizirane kvarove prikazane su na slici 6. Vrednost T60 za zdrav ET je 434 ms. Kada je prisutan najmanji kvar (slučaj SC 0-1), T60 se smanjuje na 198 ms, što je smanjenje od 55% u poređenju sa ispravnim stanjem. Vrednosti T60 se kreću od 4 ms za najteži kvar (SC 0-7) do 166.5 ms za slučaj kvara SC 0-2. Značajno smanjenje T60 prilikom postojanja kvara omogućava vrlo pouzdano otkrivanje kvara. Dalje, T60 se monotono smanjuje kako se povećava stepen kratkog spoja. Ovo daje predloženom parametru izrazitu prednost u odnosu na postojeće parametre zasnovane na SFRA u pogledu njegove sposobnosti da jasno ukaže na ozbiljnost kvara.

Upotreba zvučne kartice kao mernog uređaja umesto specijalizovanih SFRA analizatora je povoljnija pre svega u pogledu cene. Međutim, njena prednost nad SFR analizatorom je i u pogledu individualnog trajanja merenja. Trajanje merenja je oko 5 sekundi prilikom korišćenja zvučne kartice u poređenju sa 2 minuta prilikom korišćenja SFR analizatora. Sprovođenjem SFR analize i izračunavanjem predloženog indikatora (vreme reverberacije - T60) pomoću zvučne kartice može se izvršiti pouzdano otkrivanje kvarova i njihova kvantifikacija. Rezultati prikazani u ovom radu jasno ukazuju da primena predložene procedure pomoću jednostavnog audio uređaja obezbeđuje nedvosmislenu indikaciju kvarova energetskog transformatora.

V. Zaključak

Vreme reverberacije (T60) predloženo je u ovom radu kao novi parametar za detekciju kvara energetskog transformatora. Iako se pretežno koristi u akustici prostorija, vreme reverberacije (T60) u opštijem smislu predstavlja vreme potrebno da se impulsni odziv sistema smiri. Glavna prednost ovog parametra je što njegovo izračunavanje zahteva samo talasne oblike u vremenskom domenu ulaznog i izlaznog signala. Profesionalni SFRA analizatori vrše izračunavanja u frekvencijskom domenu, pa je predloženi postupak proračuna jednostavniji. Realizovani eksperimenti pokazuju sposobnost novog parametra da ukazuje na prisustvo kvara unutrašnjeg namotaja i njegovu težinu, tj. broj kratko spojenih navojaka. Pored toga, predloženi parametar ima prednost u odnosu na postojeće statističke pokazatelje zasnovane na SFRA, posebno u pogledu određivanja težine kvara. Novi indikator T60 jasno ukazuje na različite tipove grešaka bez složenih matematičkih operacija.

Za realizaciju eksperimenata korišćena je audio kartica. S obzirom na svoj jednostavniji dizajn, znatno nižu cenu i kraće vreme potrebno za merenje i obradu signala, zvučna kartica predstavlja vrlo korisnu alternativu profesionalnim SFRA analizatorima za preliminarnu dijagnostiku ET.

Budući rad biće usmeren na korišćenje vremena reverberacije za otkrivanje drugih vrsta unutrašnjih kvarova energetskog transformatora, poput deformacija namotaja. Takođe će se ispitati mogućnost korišćenja predloženog parametra za praćenje starenja energetskog transformatora.

Zahvalnica

Ovaj rad je realizovan u okviru projekta TR 36026 i projekta III 42009, koje finansira Ministarstvo prosvete, nauke i tehnološkog razvoja Republike Srbije.

Literatura

[1] S. Ghani, Y. Thayoob, Y. Ghazali, M. Khiar, I. Chairul, Distribution system performance evaluation accounting for data uncertainty, Procedia Engineering, vol.68, pp 469-476, 2013.

[2] E. Kornatowski, S. Banaszak, Diagnostics of a Transformer's Active Part with Complementary FRA and VM Measurements, IEEE Transactions on Power Delivery, vol. 29, no. 3, Berlin, pp 1398-1406, 2014.

[3] IEEE PC57.149/D8, Guide for the Application and Interpretation of Frequency Response Analysis for Oil Immersed Transformers, 2012.

[4] IEC 60076-18, Power transformers - Part 18: Measurement of frequency response,2012.

[5] CIGRE Working Group A2.26, "Mechanical-condition assessment of transformer windings using frequency response analysis", 2008.

[6] CIGRE Working Group A2,53 "Advances in the interpretation of transformer Frequency Response Analysis (FRA)", Reference 812, SeETember, 2020.

[7] J.R. Secue, E. Mombello, Sweep frequency response analysis (SFRA) for the assessment of winding displacements and deformation in power transformers, Electric Power Systems Research, vol. 78, pp. 1119–1128, 2008

[8] A. Abbasi, M. Mahmoudi, Diagnosis and Clustering of Power Transformer Winding Fault Types by Cross-Correlation and Clustering Analysis of FRA Results, IET Generation, Transmission & Distribution, vol: 12, pp. 4301-4309, 2018.

[9] A. Pandya, B. Parekh, Interpretation of Sweep Frequency Response Analysis (SFRA) trace for the open circuit and short circuit winding fault damages of the power transformer, Electrical Power and Energy Systems, vol. 62, pp. 890–896, 2014.

[10] Akshay A. Pandya, B.R. Parekh, Interpretation of Sweep Frequency Response Analysis (SFRA) traces for the open circuit and short circuit winding fault damages of the power transformer, vol. 62, pp. 890–896, 2014.

[11] Jian Sun, Qing Yang, PeiyuSu,Shilin Wu, Song Chen, Lewei He, Diagnosis of winding fault in three-winding transformer using lightning impulse voltage, Electric Power Systems Research, vol. 175, pp. 1–9, 2019.

[12] Ni Jianqiang, Zhao Zhongyong, Tan Shan, Chen Yu, Yao Chenguo, Tang Chao, The actual measurement and analysis of transformer winding deformation fault degrees by FRA using mathematical indicators, Electric Power Systems Research, vol. 184, pp. 1–11, 2020.

[13] Ali Reza Abbasi, Mohammad Reza Mahmoudi, Application of statistical control charts to discriminate transformer winding defects, Electric Power Systems Research, vol. 191, pp. 1–11, 2021.

[14] Reza Khalili Senobari, JavadSadeh, Hossein Borsi, Frequency response analysis (FRA) of transformers as a tool for fault detection and location: A review, vol. 155, pp. 172–183, 2018.

[15] Janneth R. Secue, Enrique E. Mombello, New SFRA measurement interpretation methodology for the diagnosis of power transformers, Electrical Engineering, vol. 96, pp. 183–198, 2014.

[16] Hossein Ahmadi, BehroozVahidi, Amin Foroughi Nematollahi, A simple method to detect internal and external short-circuit faults, classify and locate different internal faults in transformers, Electrical Engineering, vol. 96, pp. 183–198, 2014.

[17] M. R. Schroeder, New Method of Measuring Reverberation Time, The Journal of the Acoustical Society of America 37(6):409-412, March 1965, DOI: 10.1121/1.1909343.

[18] A Papoulis, Circuits and Systems: A Modern Approach, Oxford University Press. June 8, 1995.

[19] G.V. Stan, J. J. Embrechts, D. Archambeau, Comparison of different impulse response measurement techniques, Journal of the Audio Engineering Society 50(4):249-262, April 2002,

[20] Tehnička dokumentacija proizvođača Doble M5400 SFRA Analyzer, dostuono na: https://www.doble.com/product/m5400/, pristupano 19.6.2021.

[21] Tehnička dokumentacija proizvođača UR22 Sound Card, dostupno na http://download.steinberg.net/downloads_hardware/UR22/UR22_documentation/UR22_OperationManual_en.pdf, pristupano 19.6.2021.

[22] S. M. Al-Ameri, M. S. Kamarudin, S. Muhammad, et. al., Interpretation of Frequency Response Analysis for Fault Detection in Power Transformers, Applied Sciences, vol. 11, no. 7, 2021.

ABSTRACT

The power transformer and generator are two of the vital elements of electric power systems. The reliability of the power transformer is directly related to its condition. The analysis of the results is based on a visual comparison of the frequency responses of periodic tests of transformers, obtained with the help of specialized devices. The disadvantage of this method is the weakening of the engineering experience. An objective way to detect faults is to calculate mathematical indicators based on frequency response analysis. These parameters are often not able to unambiguously indicate the degree of failure. This paper presents a new approach for detecting internal short circuits of transformers A new indicator based on the reverberation time, commonly employed in acoustics, is proposed instead. The proposed parameter is based on time-domain calculation and enables reliable detection of transformer inter-turn short circuits and accurate estimation of their degree.

Power Transformer Reverberation Time

Miloš Bjelić, Bogdan Brković, Mileta Žarković, Tatjana Miljković

Analiza upotrebljivosti ekonomičnog audio hardvera prilikom snimanja impulsnih odziva prostorije

Marko Ličanin, Dejan Ćirić, Darko Mihajlov i Momir Praščević

***Apstrakt*—Monitoring i analiza buke u životnoj sredini, prema trenutnim standardima i metodologijama zahteva korišćenje opreme visoke preciznosti čija cena može biti izuzetno visoka. Zbog toga je monitoringom teško pokriti šira područja gde je nivoe buke potrebno pratiti u većem broju tačaka. Postavlja se pitanje da li je moguće vršiti monitoring koristeći ekonomičnija rešenja čime bi se povećao broj mernih lokacija. Prvi korak u analizi ovakvih rešenja je ispitivanje tehničkih karakteristika dostupnog hardvera koji se može iskoristiti u svrhu snimanja audio signala. Ubrzani razvoj mikro računara omogućio je njihovu integraciju u različitim projektima, gde oni služe kao centralne jedinice za obradu signala. Razvoj i unapređenje novih generacija senzora različitog tipa, imajući u vidu njihove niske cene na tržištu, omogućava praćenje velikog broja fizičkih i hemijskih veličina široj populaciji stručnjaka i entuzijasta. Istraživanje koje je ovde prezentovano odnosi se na analizu rada MEMS mikrofona kao jednog od pomenutih senzorskih uređaja, kada se on u sprezi sa Raspberry Pi mikroračunarom koristi za snimanje impulsnog odziva prostorije. Izvršeno je poređenje rezultata sa onim dobijenim mernom akustičkom opremom u istim uslovima rada.**

***Ključne reči*—Impulsni odziv, MEMS mikrofon, Latencija, Ponovljivost merenja, Raspberry Pi.**

I. Uvod

Generacije mikroračunara koje se pojavljuju poslednjih godina poseduju visoke performanse i veliki broj ulazno-izlaznih periferija. Fleksibilnost koju ovu uređaji nose u smislu programabilnosti i skalabilnosti, mala potrošnja energije i niska cena daju veliku prednost u odnosu na komercijalnu mernu opremu [1]. Jedan od tipičnih predstavnika je Raspberry Pi (RPi) [2] mikroračunar (trenutno četvrte generacije) koji ima sledeće karakteristike:

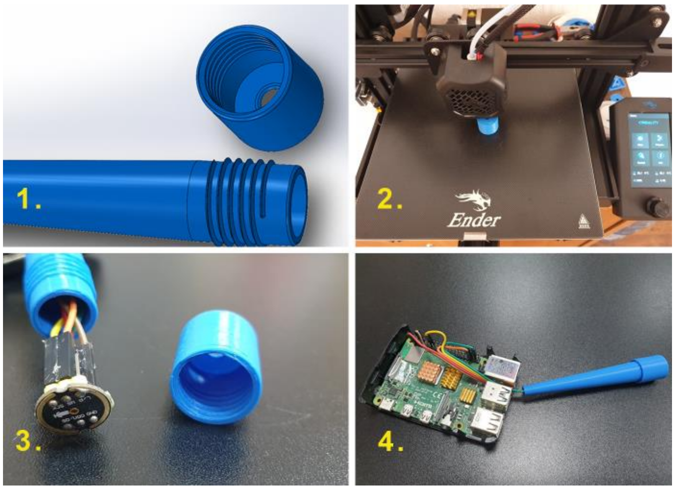
- Broadcom, Quad core Cortex-A72 (ARM v8);
- 8GB LPDDR4-3200 SDRAM;
- 2.4 GHz i 5.0 GHz wireless, Bluetooth 5.0, BLE;
- Gigabit Ethernet;
- 2 USB 3.0 porta; 2 USB 2.0 porta;
- Raspberry Pi standard 40 pin GPIO;
- 2 × micro-HDMI porta;
- 4-polni stereo audio i kompozitni video port;
- 3.5 mm analogni audio konektor.

Procesna moć ovakvog uređaja se približava performansama desktop računara uz značajno nižu potrošnju električne energije, što ga čini izuzetno primenljivim u uređajima koji se mogu napajati baterijski ili solarnom energijom. Upravo iz tog razloga, autori rada vršili su ispitivanje mogućnosti za iskorišćenje RPi pri snimanju audio signala, kao centralne jedinice za monitoring buke. Akustički senzor (mikrofon) koji je kompatibilan sa RPi je MEMS mikrofon INMP441. Vršeno je merenje impulsnog odziva prostorije korišćenjem RPi i INMP441 u dva stanja rada RPi. Snimanja su ponovljena korišćenjem merne akustičke opreme povezane na RPi i izvršeno je poređenje dobijenih rezultata. Za izdvajanje odziva upotrebljena je *swept-sine* tehnika [3][4]. Analiza je urađena tako što su određene vršne vrednosti i latencije impulsnih odziva, i rezultati su predstavljeni grafički.

II. Povezivanje opreme i proces merenja

A. Povezivanje MEMS mikrofona

INMP441 je minijaturni MEMS mikrofon montiran na okruglu PCB pločicu sa pratećim SMD komponentama. Kako bi se simulirao izgled komercijalnog mikrofona, dizajnirano je kućište u programskom paketu SolidWorks. Ovo je prikazano na Slici 1, označeno žutom bojom i brojem 1. Model kućišta je zatim odštampan na 3D štampaču Creality Ender 3 V2.



Sl. 1. Realizacija MEMS mikrona sa kućištem i povezivanje sa Raspberry Pi računarom. Brojevi označavaju korake u realizaciji.

Marko Ličanin – Univerzitet u Nišu, Fakultet zaštite na radu u Nišu, Čarnojeviča 10A, 18106 Niš, Srbija (e-mail: marko.licanin@znrfak.ni.ac.rs).
Dejan Ćirić – Univerzitet u Nišu, Elektronski fakultet u Nišu, Aleksandra Medvedeva 14 18115 Niš, Srbija (e-mail: dejan.ciric@elfak.ni.ac.rs).
Darko Mihajlov – Univerzitet u Nišu, Fakultet zaštite na radu u Nišu, Čarnojevića 10A, 18106 Niš, Srbija (e-mail: darko.mihajlov@znrfak.ni.ac.rs).
Momir Praščević – Univerzitet u Nišu Fakultet zaštite na radu u Nišu, Čarnojevića 10A, 18106 Niš, Srbija (e-mail: momir.prascevic@znrfak.ni.ac.rs).

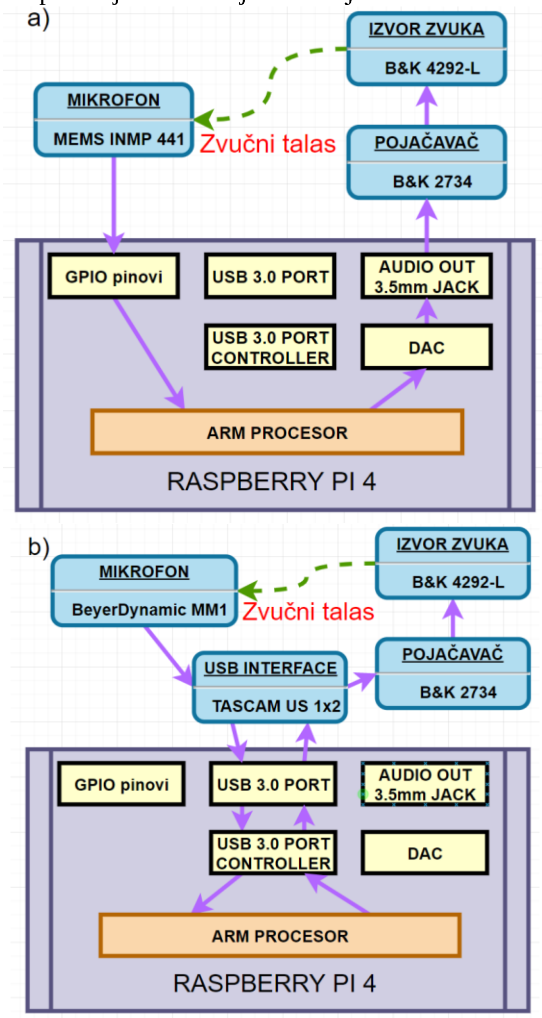
Nakon 3D štampe, montiran je mikrofon u kućište i zaštitna kapsula sa otvorom. MEMS mikrofon je zatim povezan na RPi na odgovarajuće pinove 40-pinskog konektora i u programskom jeziku Python napisana skripta za testiranje procesa snimanja.

B. Povezivanje ostale opreme

Oprema korišćena u procesu ispitivanja je sledeća :

- MEMS mikrofon InvenSense I2S INMP441;
- Merni kondenzatorski mikrofon BeyerDynamic MM1;
- USB interfejs (zvučna kartica) Tascam US2x1;
- Brüel & Kjær pojačavač tip 2734;
- Brüel & Kjær omni-direkcioni referentni izvor zvuka tip 4292-L;

Navedenom opremom realizovane su dve situacije koje su grafički prikazane na Slici 2. Razlog za to je ispitivanje uticaja različitog hardvera i putanja audio signala na latenciju impulsnog odziva prostorije, kao i na ponovljivosti latencije prilikom ponovljenih merenja. Situacije su sledeće:



Sl. 2. Grafički prikaz povezivanja mernog lanca u situaciji kada je korišćen MEMS mikrofon (a) i situaciji kada je korišćena merna oprema (b).

- Situacija 1 – MEMS mikrofon INMP 441 povezan je preko pinova sa RPi. I2S drajver uvodi signal sa mikrofona direktno u ARM procesor (Slika 2. a)). Signal reprodukcije se konvertuje preko DA konvertora i kao analogni vodi na pojačavač. Ova situacija bi trebalo da pokaže minimalnu latenciju s obzirom da je putanja kretanja signala minimalna i da nema mnogo hardvera kroz koji signal prolazi [5].
- Situacija 2 – Na RPi vezan je USB interfejs Tascam US 1x2 (zvučna kartica) preko koga se odvija snimanje i reprodukcija, između interfejsa i procesora postoji i USB 3.0 kontroler koji može uneti određeno kašnjenje, na ulaz USB interfejsa vezan je BeyerDynamic kondenzatorski mikrofon, dok je izlaz odakle se vrši reprodukcija povezan na pojačavač B&K 2734 (Slika 2. b)). Putanja kretanja signala je u ovom slučaju duža u odnosu na situaciju 1, te je očekivano da latencija ima višu vrednost [5].

C. Proces izdvajanja impulsnih odziva

Kako bi se izdvojili impulsni odzivi prostorije korišćena je tehnika snimanja pobudnim signalom promenljive sinusoide razvučene u vremenu (*swept-sine* ili *sweep* tehnika). U programskom paketu Python generisan je *swept-sine* signal i njegov inverzni filter. Signal je reprodukovan preko omni-direkcionalnog referentnog izvora zvuka B&K 4292-L i vršeno je snimanje odziva prostorije. Snimljeni signal se zatim konvoluira sa inverznim filtrom, čime se dobija impulsni odziv prostorije [6].

D. Tok merenja

Merenja su izvršena podešavajući RPi u dva stanja:

- Stanje 1 - RPi GUI je uključen;
- Stanje 2 - RPi GUI je isključen.

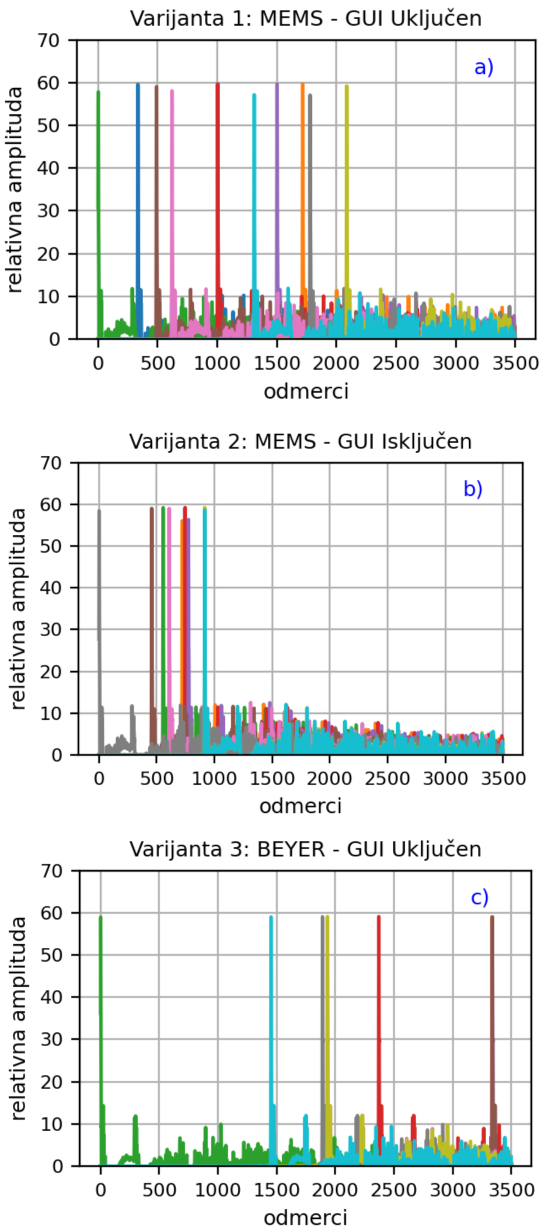
Stanje 1 podrazumeva korišćenje monitora ili udaljeno povezivanje RPi računara preko VNC servisa na pristupni terminal (u ovom slučaju laptop). Linux sistem se kontroliše kroz Desktop aplikaciju i odatle pokreće Python skripta za snimanje i reprodukciju. Stanje 2 podrazumeva ukidanje rada grafičkog interfejsa. RPi računaru se pristupa preko udaljenog terminala (laptop) preko SSH servisa i uređaj kontroliše komandama u Linux terminalu. Ova analiza izvršena je u cilju ispitivanja uticaja rada Linux grafičkog okruženja na performanse rada audio hardvera i RPi računara, prvenstveno kada se koristi MEMS mikrofon u kombinaciji sa RPi. Iz tog razloga izvršeno je snimanje impulsnih odziva u četiri varijante:

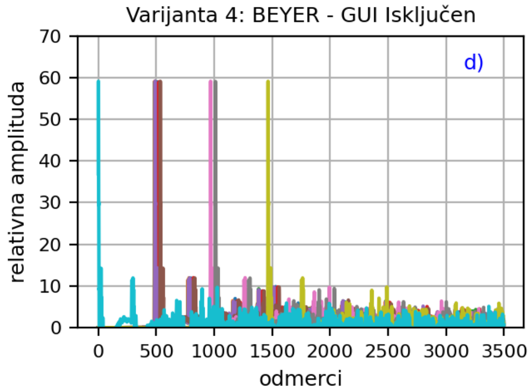
- Varijanta 1 (Situacija 1-Stanje 1): Povezan MEMS mikrofon, GUI uključen, 10 snimljenih odziva;
- Varijanta 2 (Situacija 1-Stanje 2): Povezan MEMS mikrofon, GUI isključen, 10 snimljenih odziva;
- Varijanta 3 (Situacija 2-Stanje 1): Povezan BeyerDynamic mikrofon i Tascam USB interfejs, GUI uključen, 10 snimljenih odziva;
- Varijanta 4 (Situacija 2-Stanje 2): Povezan BeyerDynamic mikrofon i Tascam USB interfejs, GUI isključen, 10 snimljenih odziva.

Koristeći interno napisane Python skripte izvršena je obrada podataka i izdvajanje impulsnih odziva, koji su zatim analizirani.

III. Rezultati

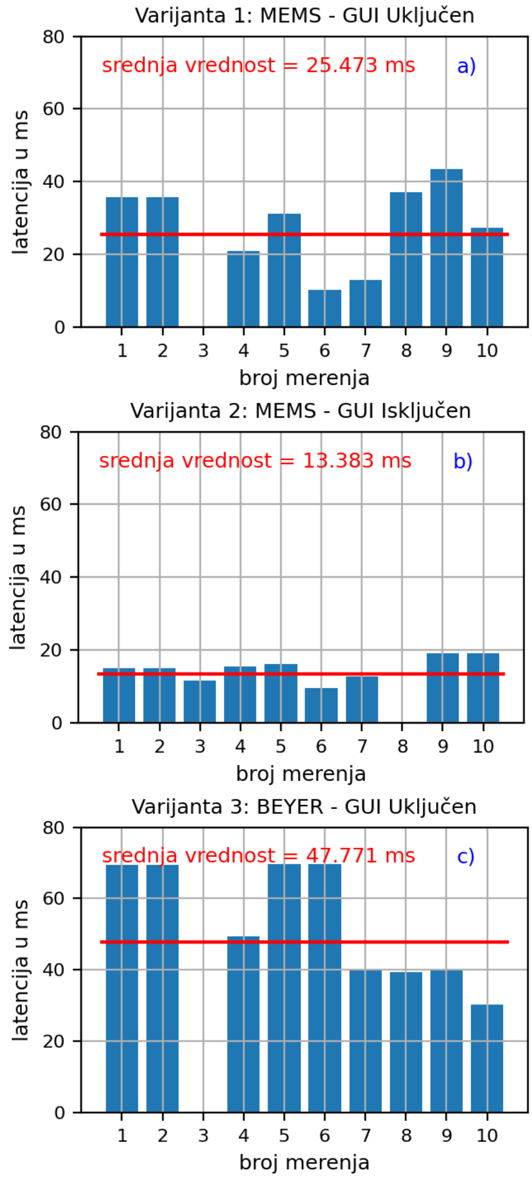
Na Slici 3 prikazani su oblici izdvojenih impulsnih odziva (kao unipolarni signali) za svaku od prethodno pomenutih varijanti i njihov pomeraj u vremenu, odnosno latencija. Latencija je određena kao relativno kašnjenje ostalih impulsnih odziva u odnosu na najmanje zakašnjeni impulsni odziv. Poređenjem varijanti 1 i 2 (Slike 3.a) i 3.b)) može se zaključiti da isključenjem grafičkog interfejsa dolazi do značajno boljeg grupisanja impulsnih odziva i smanjenja latencije. Isti zaključak je moguće doneti ukoliko se uporede rezultati kod varijanti 3 i 4 (Slike 3.c) i 3.d)). Kada se koristi merna oprema umesto MEMS mikrofona, takođe dolazi do smanjenja latencije isključenjem grafičkog interfejsa na RPi. Poređenjem situacija 1 i 2 kada je u pitanju latencija, može se zaključiti da ona ima manju vrednost u korist situacije 1 (poređenje slika 3.a) sa 3.c) i 3.b sa 3.d)

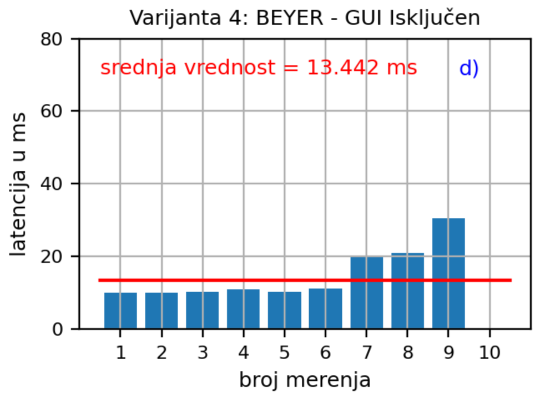




Sl. 3. Raspored impulsnih odziva predstavljen u odnosu na najmanje zakašnjeni odziv kod koga je latencija 0. Rezultati su dati za različite varijante merenja (varijante 1-4 odgovaraju slikama a)-d) respektivno).

Analiza ponovljivosti latencije može se bolje uvideti na Slici 4. Primetno je da srednja vrednost latencije impulsnih odziva opada značajno isključenjem grafičkog interfejsa RPi (poređenje slika 4.a) sa 4.b) i 4.c sa 4.d).

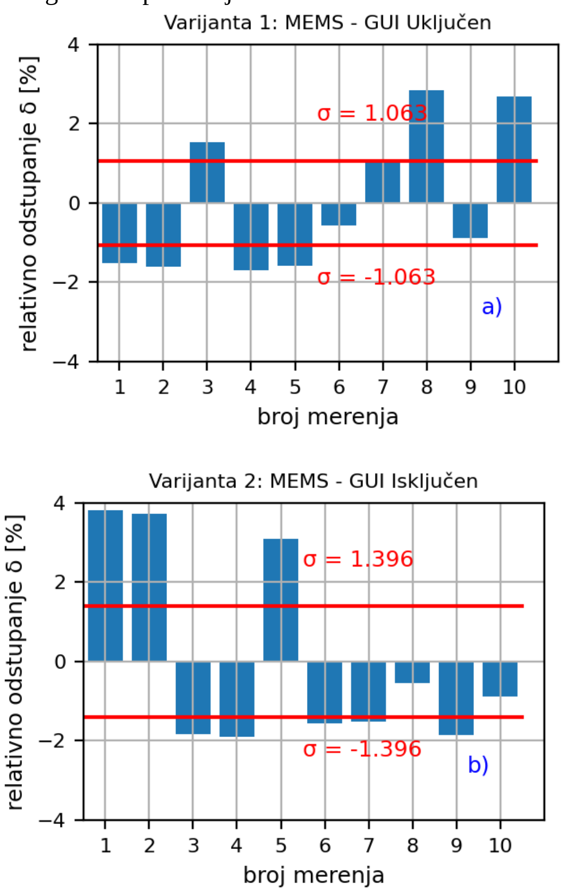


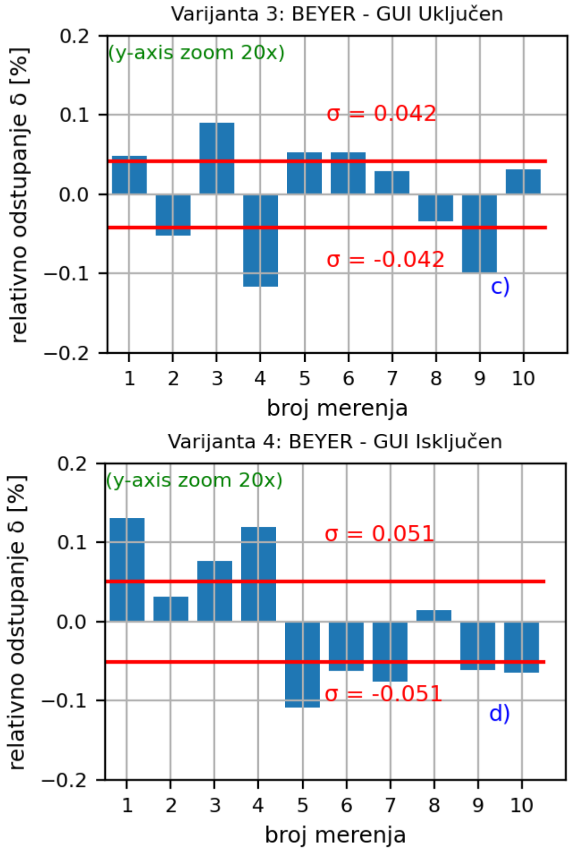


Sl. 4. Latencija impulsnih odziva u milisekundama predstavljena u odnosu na najmanje zakašnjeni odziv. Rezultati su dati za različite varijante merenja (varijante 1-4 odgovaraju slikama a)-d) respektivno).

Posmatranjem varijante 2 na Slici 4.b) može se utvrditi da je ponovljivost latencije izuzetno dobra za slučaj kada je grafički interfejs isključen. To dovodi do zaključka da sa stanovišta ponovljivosti latencije, sasvim je validno koristiti MEMS mikrofon u kombinaciji sa RPi, kao i da je neophodno isključiti grafički interfejs mikroračunara RPi.

Poslednja analiza odnosi se na ponovljivost amplituda impulsnog odziva prostorije. Kao i u prethodnim slučajevima, vršeno je merenje u četiri varijante, što je prikazano na Slici 5. Na *y* osama bar grafika prikazano je relativno odstupanje pojedinačnih amplituda od srednje proračunate vrednosti za 10 merenja. Na *x* osama obeležen je redni broj merenja impulsnog odziva prostorije.





Sl. 5. Relativno odstupanje amplituda impulsnih odziva od srednje vrednosti proračunate za 10 merenja. Rezultati su dati za različite varijante merenja (varijante 1-4 odgovaraju slikama a)-d) respektivno).

Posmatrajući varijante 3 i 4 (Slika 5.c) i 5.d)), i poredeći rezultate sa varijantama 1 i 2, (Slika 5.a) i 5.b)) može se zaključiti da je ponovljivost amplituda značajno bolja kada je u pitanju merna oprema u odnosu na MEMS mikrofon. *X* – osa je u slučaju prezentovanih rezultata kod merne opreme uveličana 20 puta kako bi se uočila razlika u amplitudama. Takođe se može zaključiti da promena stanja kod RPi (uključenje ili isključenje grafičkog interfejsa) ne utiče na promenu ponovljivosti amplitude impulsnih odziva.

IV. Zaključak

Izvršena ispitivanja predstavljaju samo jedan od koraka u evaluaciji ograničenja i prednosti koje pruža ekonomični hardver pri korišćenju u audio aplikacijama. Vršena je analiza ponovljivosti latencije i amplitude impulsnih odziva prostorije u četiri varijante. Sagledavan je uticaj grafičkog interfejsa na latenciju i amplitudu impulsnih odziva. Analiza je urađena za slučaj kada je MEMS mikrofon povezan na Raspberry Pi 4, kao i za slučaj kada je kao audio interfejs korišćena USB kartica sa mernim mikrofonom.

Ispitivanje je pokazalo da postoji značajan uticaj grafičkog interfejsa na ponovljivost latencije čime je utvrđeno da je bolje raditi samo sa Linux terminalom kada je audio snimanje u pitanju. Pokazano je MEMS mikrofoni imaju ograničenja po pitanju amplitude kada se rezultati ponovljivosti amplituda porede sa ponovljivošću kod studijske opreme.

U budućnosti treba izvršiti dalja ispitivanja, koja se odnose na ponovljivost frekvencijskih karakteristika impulsnih odziva, kao i ovde prezentovane rezultate potvrditi novim proširenim setom merenja.

ZAHVALNICA

Ovaj rad je podržan od strane Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije. Deo opreme koji je iskorišćen u realizaciji rada nabavljen je u okviru Erasmus + projekta SENVIBE (ref. br. 598241-EPP-1-2018-1-RS-EPPKA2-CBHE-JP) koji je finansiran od strane Evropske unije.

LITERATURA

[1] M. Ličanin, M. Praščević, D. Mihajlov "Realization of the low-cost noise measurement monitoring station using MEMS microphone technology and micro-PC, "In: Proc. of 26th International Conference: Noise and Vibration, pp. 133–136, Niš, Srbija: Fakultet zaštite na radu u Nišu, 2018.

[2] D. Molloy, *Exploring Raspberry Pi: Interfacing to the Real World with Embedded Linux, 1*st ed. Springdale, Arkansas, US: Willey, June 13, 2016.

[3] S. Müller S. P. Massarani, "Transfer-Function Measurement with Sweeps" J. Audio Eng. Soc., vol. 49, no. 6, pp. 443–471, 2001

[4] S. Müller, "Measuring transfer-functions and impulse responses," in *Part I: Acoustic Signals and Systems,* in *Handbook of Signal Processing in Acoustics*, pp. 65–86, New York, USA: Springer Science, 2008

[5] M. Jelenković, D. Ćirić, M. Ličanin, A. Đorđević "Ponovljivost impulsnog odziva prostorije," In: Zbornik radova konferencije ETRAN, Društvo za ETRAN, Banja Vrućica, Srbija, pp. AK.2.7-1–AK.2.7-4., 2011.

[6] A. Farina, "Simultaneous Measurement of Impulse Response and Distortion with a Swept-sine Technique," 108th Convention Audio Eng. Soc., abstract in J. Audio Eng. Soc., vol. 48, no. 4, p. 350, 2000.

ABSTRACT

Monitoring and analysis of environmental noise, according to current standards and methodology, requires the use of high-precision equipment whose price can be extremely high. Therefore, it is difficult to cover wider areas where noise levels need to be monitored at a large number of points. The question is if it is possible to perform the level supervision with low-cost solutions, which increases the number of measuring locations. The first step in the analysis of such solutions is to examine the technical characteristics of the available hardware that can be used for the purpose of recording audio signals. The accelerated development of microcomputers enables their integration into various projects, where they serve as central signal processing units. The development and improvement of new generations of sensors of different types, having in mind their low prices on the market, enables the monitoring of a large number of physical and chemical quantities by a wide population of experts and enthusiasts. Research presented here refers to the analysis of the operation of the MEMS microphone as one of the mentioned sensor devices, when it is used in conjunction with the Raspberry Pi microcomputer to record the room impulse responses. A comparison of the results with those obtained using the low-cost measurement acoustic equipment in the same working conditions have been done here.

Analysis of the usability of economical audio hardware when recording room impulse responses

Marko Ličanin, Dejan Ćirić, Darko Mihajlov, Momir Praščević

Uticaj COVID 19 zaštitnih maski na razumljivost govora u srpskom jeziku

Miloš Bjelić, Tatjana Miljković, Miomir Mijić, Dragana Šumarac Pavlović

***Apstrakt*—U ovom radu prikazana je analiza uticaja zaštitnih maski za lice na razumljivost govora. Analizirana su tri različita tipa zaštitnih maski koje se koriste u pandemijskim uslovima (pamučna maska, hirurška maska i maska N95). Takođe, analiziran je uticaj zaštitnog transparentnog vizira. Na osnovu govora nekoliko osoba utvrđen je oblik spektra dugovremenog govora u srpskom jeziku u slučaju upotrebe zaštitnih sredstava. Izvršeno je nekoliko subjektivnih testova u kojima je merena logatomska razumljivost govora sa i bez zaštitnih maski. Eksperimenti su organizovani u kontrolisanim uslovima (slušanje preko slušalica) i u prostoriji sa velikim vremenom reverberacije. Pokazano je da se prilikom korišćenja zaštitne maske N95 ostvaruje najbolja razumljivost govora u srpskom jeziku u odnosu na druga dva tipa analiziranih maski.**

***Ključne reči* — Covid19, logatomi, N95, razumljivost govora, slabljenje, subjektivni testovi, zaštitne maske.**

I. Uvod

Prijem govornih informacija je proces prepoznavanja pojedinačnih zvučnih simbola, ali i čitavih reči i rečenica kao celina. Razumljivost govora pri prenosu kroz neki fizički kanal smanjivaće se kada neki parazitski signali prekrivaju delove govornog signala manjih amplituda. Vokali su delovi govornog signala koji imaju relativno veliku energiju dok su konsonanti delovi koji su značajno slabiji. Prekrivanje tiših delova govornog signala može nastati na dva načina: kao vremenski uniformno pokrivanje aditivnim šumom i kao vremenski ograničeno pokrivanje refleksijama u prostoriji koje stižu sa kašnjenjem nakon jačih delova govornog signala (vokala) [1]. Reverberacioni proces u značajnoj meri može smanjiti razumljivost govora. Pokazano je da se razumljivost govora manja za 15% u prostorijama sa velikim vremenom reverberacije (T60 > 2.5 s) u odnosu na akustički obrađene prostore [2]. U literaturi je pokazano da razumljivost govora pri malim odnosima signal-šum, npr. manjim od 5 dB, može biti ispod 75% [3-4].

U poslednje skoro 2 godine pandemija virusa Covid19 uticala je na život ljudi na celom svetu. Preduzimane su mere predostrožnosti kako bi se smanjilo širenje virusa. Pored socijalne distance i karantina, upotrebu sredstava za dezinfekciju sprovede se mere koje su podrazumevale obaveznu upotrebu zaštitne opreme, kao što su zaštitne maske, viziri za lice i rukavice. Upotreba zaštitnih maski, u velikoj meri, sprečava širenje virusa. Ove vrste zaštite pored sprečavanja širenja virusa imaju uticaj i na komunikaciju između ljudi. Maske za lice su vidljiva prepreka i utiču na verbalnu komunikaciju, emocionalna ekspresiju i čitanje sa usana, što je vrlo korisno pomagalo u uobičajenoj komunikaciji među ljudima i izuzetno korisno za ljude sa oštećenim sluhom [5].

Zbog navedenih razloga veliki broj istraživača u oblasti akustike bavio se uticajem zaštitnih maski na govornu komunikaciju. Istraživanja su se pre svega bavila uticajem zaštitnih maski na razumljivost govora [2-3], [5-12]. Analizirana je razumljivost u prostorijama, kao što su školske učionice [6]. Pokazano je da zaštitne maske utiču na zamaranje slušalaca prilikom dužeg slušanja, kao posledica smanjene razumljivosti [7]. Analiziran je i uticaj zaštitni maski na govor prilikom reprodukcije govora preko slušalica [3]. Takođe, ispitivan je uticaj zaštitne opreme na usmerenost govornika [8]. Posebno su analizirani uticaji maske na oblike dugovremenog spektra govora [3], [9-10], kao objektivna mera uticaja zaštitne opreme. U nekim istraživanjima posebno je ispitivan uticaj transparentnog zaštitnog vizora na govornu komunikaciju [9, 10]. Često su korišćeni subjektivni testovi kao metod za ispitivanje nekih od hipoteza vezanih za zaštitne maske [11].

Istraživanja koja se mogu pronaći u literaturi se pre svega odnose na ispitivanje uticaja maski na govor u engleskom jeziku. Ideja istraživanja u ovom radu bila je da se sagleda uticaj zaštitnih maski i vizira na komunikaciju u srpskom jeziku. Analizirana je promena oblika dugovremenog spektra govornog signala prilikom korišćenja različitih tipova zaštitnih maski. Ovi eksperimenti imali su za cilj određivanje objektivnog uticaja maski. Sprovedeni su eksperimenti u kojima je određivana razumljivost govora na srpskom jeziku. Razumljivost govora određivana je za slučaj idealnog slušanja govora i slušanje u realnim uslovima u prostoriji. Da bi se odredila razumljivost govora korišćeni su subjektivni testovi.

Rad je organizovan u pet poglavlja. U drugom poglavlju predstavljena je metodologija. U narednom poglavlju prikazana je eksperimentalna postavka. U četvrtom poglavlju prikazani su dobijeni rezultati i njihova diskusija. Na kraju izneti su zaključci na osnovu sprovedenog istraživanja.

Miloš Bjelić – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: bjelic@etf.rs).

Tatjana Miljković – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: tm@etf.rs).

Miomir Mijić – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: emijic@etf.rs).

Dragana Šumarac Pavlović – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: dsumarac@etf.rs).

II. METODOLOGIJA

Prvi deo istraživanja u ovom radu obuhvata analizu oblika spektra dugovremenog govora. Oblik spektra normalnog govora na srpskom jeziku može se pronaći u literaturi [13], međutim oblik spektra govora sa zaštitnim sredstvima ne postoji. Za određivanje oblika spektra potrebno je snimanje većeg broja govornika koji izgovaraju neutralni govor. Govor je potrebno snimiti kada govornici nose zaštitne maske, odnosno vizir i u slučaju kada govornik ne nosi nikakvu zaštitnu opremu (normalan govor). Ovaj eksperiment predstavlja objektivni način za sagledavanje uticaja zaštitne opreme u pandemijskim uslovima na govornu komunikaciju.

Drugi deo istraživanja predstavlja realizaciju subjektivnih testova u kojima su ljudi ocenjivali razumljivost govora prilikom korišćenja zaštitne opreme. U slučaju izgovora reči bez smisla prepoznavanje izgovorenog teksta je moguće samo u okolnostima kada je omogućeno prepoznavanje svakog pojedinačnog glasa, a za to je potreban dovoljno visok nivo kvaliteta prenosa [14]. Iz toga sledi da samo reči bez smisla mogu dati izvesnu objektivnost u oceni kvaliteta prenosa kroz kanal bez uticaja sofisticiranih mehanizama percepcije koji nadoknađuju nedostatke u govornom signalu. Metoda koja se koristi merenje kvaliteta kanala za prenos govora i kvantifikovanje razumljivosti zasnovana je na emitovanju posebno sastavljenih reči bez smisla koje se nazivaju logatomi [14]. Proces merenje se zasniva na emitovanju logatoma na ulazu u prenosni kanal i beleženju onoga što se čuje na mestu prijema. Utvrđuje se procenat tačno primljenih logatoma i taj podatak predstavlja ocenu kvaliteta prenosa govora koji se naziva „logatomska razumljivost".

Logatomi se posebno dizajniraju za potrebe merenja razumljivosti, i obično se sastoje od tri ili četiri glasa u redosledu konsonant–vokal-konsonant (CVC) ili konsonant–vokal–konsonant-vokal (CVCV) [1, 14]. Logatomi se prave u grupama od po 50 ili 100. Pri tome se grupe tako formiraju da budu fonetski balansirane (PB word–*Phonetically Balanced word*), što znači da je u jednoj grupi logatoma statistika zastupljenosti pojedinih glasova ista kao u običnom govoru.

Logatomska razumljivost je apsolutni pokazatelj kvaliteta prenosa govora jer se dobija neposrednim slušanjem i zbog toga je u ovom radu izabrana za ispitivanje uticaja zaštitnih maski na razumljivost govora. Sprovođenjem subjektivnih testova u kojima se određuje logatomska razumljivost mogu se odrediti uticaji različitih zaštitnih maski i vizira na govornu komunikaciju.

III. POSTAVKA EKSPERIMENATA

U eksperimentima su korišćena 3 tipa zaštitnih maski: pamučna maska, hirurška maska, maska N95 i zaštitni vizir. Na slici 1. prikazani su izgledi korišćene zaštitne opreme. Analiziran je uticaj zaštitnog vizira koji se često koristiti u situacijama gde je izloženost virusu velika, kao što su bolnice. Zaštitni vizir se koristi u kombinaciji sa zaštitnim maskama, pa je u ovom radu razmatrana upotreba u kompletu sa zaštitnom hirurškom maskom. Izgled korišćenog vizira dat je na slici 1 d).

Sl. 1. Zaštitna oprema korišćena u eksperimentu: a) hirurška maska, b) N95 maska, c) pamučna maska d) zaštitni vizir

Snimanja govora sa i bez zaštitne opreme obavljena su u anehoičnoj prostoriji („gluva soba"). Govornici su se nalazili 20 centimetara ispred mikrofona i izgovarali potrebne reči, odnosno tekst. Na slici 2 prikazana je postavka snimanja govora u anehoičnim uslovima.

Sl. 2. Snimanje govora sa zaštitnim maskama u gluvoj sobi

U prvom eksperimentu snimani su govornici koji su čitali isti književni tekst u trajanju od 150 sekundi. U eksperimentu je učestvovalo 10 osoba, 5 muških i 5 ženskih. Svi govornici su izgovarali tekst bez maske, a zatim sa svim tipovima maske korišćenih u ovom radu. To znači da je svaki govornik izgovarao je isti tekst 5 puta. Na osnovu snimljenih signala formirane su grupe snimaka prema tipu maske i računati dugovremeni 1/3 oktavni spektri („tercni" spektri) [15]. Cilj ovog eksperimenta bio je sagledavanje uticaja zaštitne opreme na oblik spektra govora. Za snimanje je korišćen neusmereni mikrofon [16] i AD konvertor sa 24 bita [17].

Da bi se sagledao subjektivni uticaj zaštitne opreme na razumljivost govora bilo je neophodno sprovesti subjektivne testove. U anehoičnim uslovima snimljeno je 5 grupa logatoma od po 50 reči. Svaka od grupa predstavlja jedan od slučajeva koji se analizira (jedan bez maske i 4 sa zaštitnom opremom). Sprovedena su 2 subjektivna testa. Prvi test izveden je u kontrolisanim uslovima, tj. slušanjem snimljenih grupa logatoma preko slušalica. U ovom eksperimentu eliminisan je uticaj spoljašnje sredine, odnosno prostorije i ambijentalne buke. U testu je učestvovalo 12 osoba koje su slušale grupe logatoma preko istih slušalica [18].

Sl. 3. Izgled zvučnika korišćenog za reprodukciju logatoma u učionici

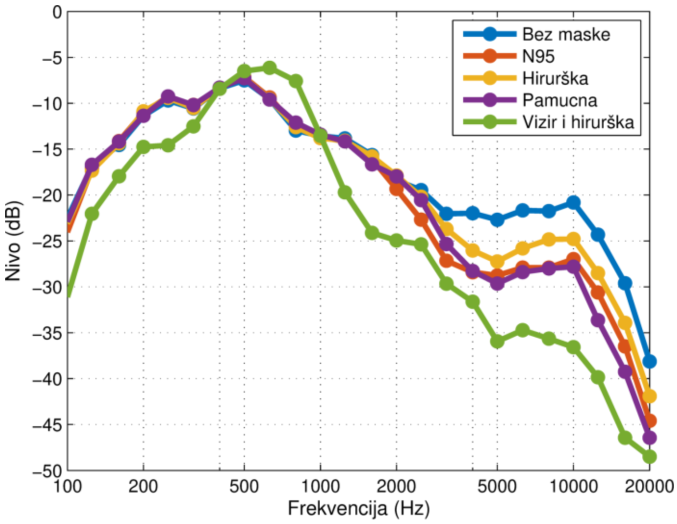
Drugi test realizovan je u prostoriji koja je imala relativno veliku vrednost vremena reverberacije. Dodatno, u prostoriji je postojala ambijentalna buka zbog loše izolacione moći vrata i prozora. Korišćena prostorija je učionica na fakultetu. Grupe logatoma su reprodukovane slušaocima pomoću usmerenog zvučnika koji je prikazan na slici 3. Pozicija zvučnika odgovarala je uobičajenoj poziciji predavača u učionici – na bini ispred table. U eksperimentu je učestvovalo 14 slušalaca, koji su raspoređeni uniformno po prostoriji, kao na slici 4. U oba subjektivna testa slušaoci su pripadali starosnoj grupi od 20 do 30 godina, što predstavlja najbolju starosnu grupu za subjektivna audio ispitivanja.

Sl. 4. Realizacija subjektivnog testa u učionici

IV. Rezultati i diskusija

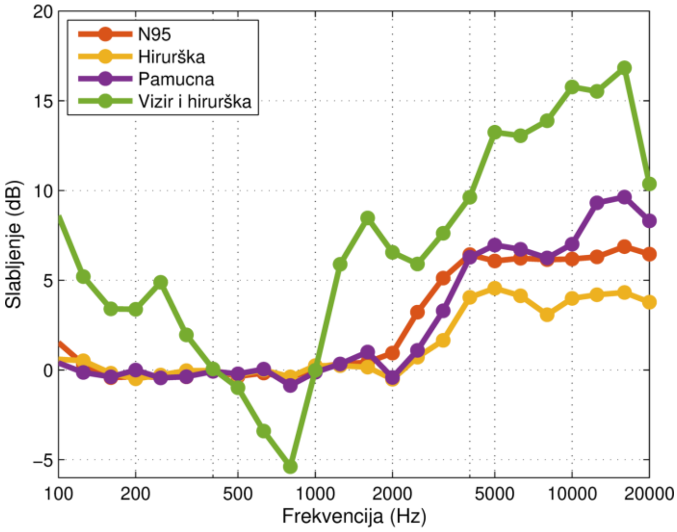
A. *Dugovremeni spektri govora*

Na slici 5 prikazani su usrednjeni dugovremeni spektri za 10 govornika. Spektri su normalizovani tako da ukupan nivo za svaki od spektara iznosi 0 dB. Plavom bojom prikazan je usrednjeni spektar govora bez korišćenja zaštitnih maski, odnosno zaštitne opreme. Maksimalna vrednost nivoa dugovremenog govora odgovara 1/3 opsegu 500 Hz. U oblasti frekvencija ispod 500 Hz spektar opada 10 dB po oktavi, a u oblasti do 3100 Hz spektar govora opada sa 6 dB po oktavi. U opsegu frekvencija između 3100 Hz i 10 kHz spektar je ravan, a nakon 10 kHz se može smatrati da nema značajnih komponenti u spektru dugovremenog govora. Ovakav oblik spektra poklapa se sa podacima o srpskom jeziku koji se mogu pronaći u literaturi [13].



Sl. 5. Izgled 1/3 oktavnih dugovremenih spektara

Dugovremeni spektar govora u slučaju korišćenja zaštitnih maski (hirurška, pamučna i N95) u oblasti frekvencija do 2000 Hz poklapa se sa oblikom spektra za govor bez maske. Prilikom korišćenja hirurške maske nivo spektar govora je za oko 5 dB manji od spektra nivoa govora bez maske, za oblast frekvencija iznad 2 kHz. Maska N95 unosi veće slabljenje u odnosu na hiruršku masku u oblasti visokih frekvencija. Vrednosti slabljenja u odnosu na normalan govor iznosi oko 6 dB. Pamučna maska unosi najveće slabljenje u odnosu na ostala dva tipa maski. Slabljenje u odnosu na govor bez maske iznosi oko 7 dB. Na osnovu vrednosti slabljenja koje unose zaštitne maske može se zaključiti da one za govor deluju kao filtar propusnih niskih frekvencija. Očekivano bi bilo da je razumljivost govora obrnuto srazmerna vrednosti slabljenja koje unosi maska. Zbog toga bi razumljivost govora trebala biti najveća za hiruršku masku, zatim za N95 i najmanja za pamučnu masku.



Sl. 6. Slabljenje zaštitnih maski izračunato na osnovu dugovremenih spektara

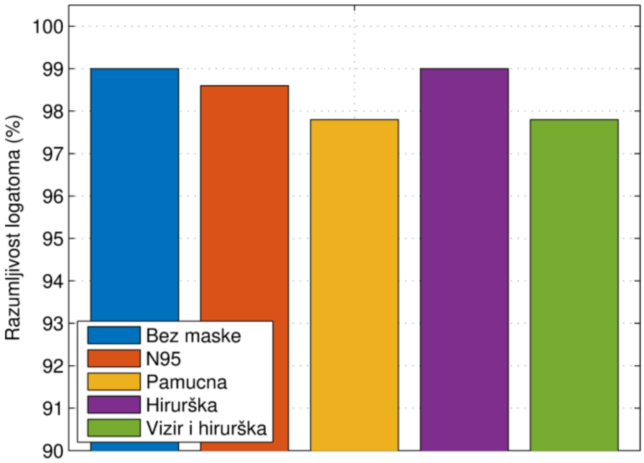
U slučaju korišćenja zaštitnog vizira (sa hirurškom maskom) oblik spektra je značajno drugačiji od spektra govora bez i sa korišćenjem zasitnih maski. Kako bi se lakše sagledao uticaj zaštitnog vizira na spektar govora izračunato je slabljenje koje unosi zaštitna oprema. Slabljenje je izračunato tako što je od spektra dugovremenog govora bez zaštitnih maski oduzet spektar govora sa maskama (Slika 5). Na slici 6 prikazano je slabljenje koje unose zaštita oprema (maske i vizir). Sa slike se mogu uočiti razlike u vrednostima slabljenja između pojedinačnih maski koje su prikazane i na slici 5.

U obliku spektra dugovremenog govora govornika koji su nosili zaštitni vizir može se uočiti zona između 400 Hz i 1000 Hz u kojoj je koncentrisana najveća energija govornog signala. Opadanje nivoa ka nižim frekvencijama je slično kao kod govora sa zaštitnim maskama, a opadanje ka višim frekvencijama iznosi oko 12 dB po oktavi. Na osnovu oblika spektra zaključuje se da zaštitni vizir deluje kao filtar propusnik opsega. Hirurška maska koja je korišćena zajedno sa vizirom doprinosi samo dodatnom slabljenju na visokim frekvencijama. Govor prilikom korišćenja vizira zvuči „telefonski“ zbog efekta filtra propusnika opsega. U literaturi [9, 10] je pokazano da se ovakav oblik spektra javlja prilikom korišćenja zaštitnog vizira i za druge jezike. Međutim, nije dato objašnjenje zašto nastaje ovaj lokalni maksimum u spektru govora. Pokazano je da se takav oblik javlja u eksperimentima u kojima je izvor zvuka čovek [6, 9, 10], ali i veštački glas [10]. To znači da artikulacija kod čoveka prilikom izgovora nije odgovorna za nastanak ovog fenomena. Fenomen koji se javlja je posledica nastanka rezonance u vazdušnom prostoru između lica i zaštitnog vizira, po kraćoj dimenziji vizira. Zbog slabljenja koje unosi vizir na visokim frekvencijama očekivano je da razumljivost govora u ovom slučaju bude najmanja.

B. Logatomska razumljivost – slušalice

Za svaku od zaštitnih maski (zaštitne opreme) pripremljena je različita grupa logatoma. Na slici 7 prikazani su rezultati logatomske razumljivosti za subjektivni test u kome su slušaoci slušali logatome preko slušalica. Rezultati su usrednjeni za svaku od grupa logatoma na osnovu rezultata dobijenih za 12 slušalaca za tu grupu. Slušanje preko slušalica predstavlja praktično idealan način slušanja po pitanju ambijentalne buke, tj. odnosa signal šum. Nivo šuma određen je nivoom šuma u gluvoj sobi gde su snimljene grupe logatoma. Odnos signal šum prilikom reprodukcije logatoma preko slušalica u ovo eksperimentu iznosio je 40 dB.

Logatomska razumljivost za slučaj govora bez zaštitne maske iznosi 99%, što predstavlja jako dobru razumljivost. To znači da slušaoci u proseku nisu razumeli samo jednu reč od 50. Razumljivost za grupu logatoma koja je snimljena kada je govornik nosio hiruršku zaštitnu masku iznosi takođe 99%. Na osnovu toga može se zaključiti da ovaj tip zaštitne maske ne kvari razumljivost govora. U slučaju korišćenja maske N95 logatomska razumljivost iznosi 98.6%, što takođe predstavlja jako dobru razumljivost govora. Nešto manja logatomska razumljivost dobijena je za slučaj korišćenja dvoslojne pamučne maske i u slučaju korišćenja zaštitnog vizira sa hirurškom maskom. Razumljivost u ovim slučajevima je ista i iznosi 97.8%.



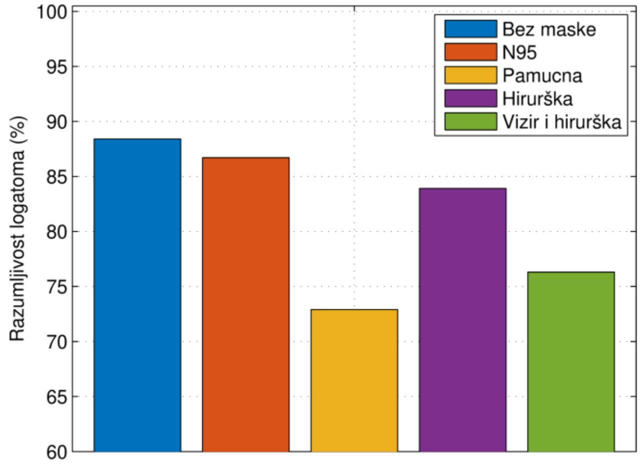
Sl. 7. Logatomska razumljivost za slušanje preko slušalica

Razlike u razumljivosti dobijene za različite tipove zaštitne opreme su relativno male pre svega zbog velikog odnosa signal šum, odnosno idealnih uslova slušanja. Vrednosti su nešto manje za pamučnu masku i zaštitni vizir, ali u ovim slučajevima slušaoci, u proseku, nisu razumeli samo dve reči. Rezultati dobijeni u ovom eksperimenti mogu pokazati koje grupe maski mogu značajnije uticati na razumljivost govora, ali su procenti razumljivosti jako slični pa se neki značajniji zaključci ne mogu napraviti. Dodatno, na osnovu rezultata u ovom subjektivnom testu ne može se sagledati uticaj zaštitnih maski na razumljivost govora u svakodnevnim uslovima. U svakodnevnom životu nivo ambijentalne buke je veći, a često i značajno veći, od nivoa ambijentalne buke u ovom eksperimentu. Ambijentalna buka može maskirati delove govornog signala, što je značajno pre svega na višim frekvencijama gde zaštitne maske dodatno slabe nivo govornog signala.

C. Logatomska razumljivost – učionica

U ovom subjektivnom testu 14 slušalaca slušalo je u učionici površine 100 m^2 grupe logatoma, koji su snimljeni sa različitim zaštitnim maskama. Učionica nije akustički obrađena, što se može videti na slici 4. Vreme reverberacije iznosi oko 2 sekunde za 1/1 oktavne frekvencijske opsege sa centralnim frekvencijama 250 Hz, 500 Hz, 1000 Hz i 2000 Hz. Nivo ambijentalne buke u prostoriji iznosio je 36.5 dBA. Merenje ambijentalne buke izvršeno je u toku trajanja eksperimenta, u sredini prostorije. Odnos signal šum prilikom reprodukcije logatoma iznosio je maksimalno 11 dB. Na slici 8 prikazani su rezultati logatomske razumljivosti u učionici za različite tipove zaštitnih maski. Rezultati pojedinačnih slušalaca su usrednjeni za odgovarajuće grupe logatoma.

Vrednost logatomske razumljivosti u prostoriji za slučaj bez korišćenja zaštitne opreme iznosi 88.4%. Vrednost razumljivosti u ovom eksperimentu je za oko 10% manja nego u slučaju eksperimenta sa slušalicama. Najbolja razumljivost ostvarena je kod slušalaca koji su sedeli u prednjoj zoni učuonice i maksimalno je iznosila 94%. Za slušaoce koji su sedeli u zadnjem delu sale razumljivost je značajno manja i minimalno iznosi 82%. U slučaju korišćenja zaštitne opreme najbolja razumljivost ostvarena je za masku N95. Logatomska razumljivost za masku N95 iznosi 86.7%, što je približno vrednosti koja je ostvarena bez korišćenja maski. Kao i u prethodnom slučaju slušaoci koji su sedeli bliže zvučniku imali su veću uspešnost prilikom prepoznavanja logatoma. Vrednosti razumljivosti za ovaj tip maske u prostoriji kreću se u opsegu od 74% do 94%.



Sl. 8. Logatomska razumljivost za slušanje u Sali

Za slučaj u kom je korišćena hirurška zaštitna maska logatomska razumljivost je nešto manja u odnosu na prethodna dva slučaja i iznosi 83.9%. Raspodela razumljivosti po prostoriji za hiruršku masku ne poklapa se sa raspodelom sa slučajem maske N95 i slučajem bez korišćenja maske. U ovom slučaju vrednosti razumljivosti su potpuno slučajne u odnosu na rastojanje slušalaca od zvučnika preko koga su emitovani logatomi. Ostvarena razumljivost govora u slučaju u kom je korišćen zaštitni vizir zajedno sa hirurškom maskom iznosi 76.3%. Ova vrednost je za 22% manja u odnosu na slučaj razumljivosti govora bez korišćenja zaštitne maske. Vrednost razumljivost je očekivano manja u odnosu na slučaj kada je korišćena samo hirurška zaštitna maska. Vrednost je manja za 7.6%, što se može pripisati uticaju vizira. Raspodela razumljivosti je slučajna po prostoriji, i ne zavisi od rastojanja slušaoca od zvučnika. Najmanja razumljivost od svih analiziranih slučajeva postignuta je u slučaju korišćenja dvoslojne pamučne maske i iznosi samo 72.9%. Ova vrednost razumljivosti je manja za čak 25.5% u odnosu na razumljivost govora ostvarenu za slučaj bez zaštitne opreme.

Neke od grešaka koje su pravili slušaoci su nastajale zamenom parova /f/-/h/, /f/-/p/ i /d/-/g/. Fonemi /g/ i /d/ pripadaju grupi ploziva i u artikulacionom smislu su slični jer se sastoje iz dva dela okluzije i eksplozije. Kod fonema /d/ eksplozija stvara šumnu energiju koja se karakteriše sa više koncentrata gotovo ravnomerno do 8 kHz [1]. Prilikom izgovora fonema /g/ pojavljuju se 2 koncentrata energije, veći na 2000 Hz i manji na oko 5 kHz. Zbog slabljenja zaštitnih maski iznad 2000 Hz slušaoci slovo /d/ mešaju sa slovom /g/. Fonem /f/ pripada grupi frikativa. Energija kod šumnog fonema /f/ se prostire od 500 Hz do gotovo 10 kHz [1]. Zbog uticaja maski u ovom frekvencijskom području razumljivost ovog fonema može biti narušena.

Posmatrajući oblik usrednjenih spektara sa slike 5 zaključuje se da maska N95 unosi veće slabljenje u odnosu na hiruršku masku. Međutim, na osnovu rezultata subjektivnog testa realizovanog u realnim uslovima (npr. slušanje predavanja u učionici) zaključuje se da je najmanji uticaj na razumljivost govora ima maska N95. Maska N95 skoro da ima isti uticaj na spektar dugovremenog govorna kao i pamučna maska, ali je razumljivost govora u učionici prilikom korišćenja pamučne maske za 14% manja u odnosu na masku N95. Takođe, slabljenje koje unosi zaštitni vizir sa hirurškom maskom je značajno veće u odnosu na sve tipove maski, ali je razumljivost govora u tom slučaju veća od razumljivosti za pamučnu masku. Na osnovu toga zaključuje se da posmatranje maske (odnosno zaštitne opreme) kao filtra za govorni signal nije dobar način za sagledavanje uticaja maski na razumljivost govora. Ponašanje maske kao filtra propusnika niskih frekvencija nije zanemarljiv uticaj na razumljivost, ali nije jedini. U literaturi se mogu pronaći podaci i o drugim uticajima maske na govor [13], pored filtriranja. Pored unošenja slabljenja na visokim frekvencijama maske utiču i na način izgovora slova, odnosno na artikulaciju. Pamučna maska je pripijena uz lice i postoji konstantno trenje usana o masku. Prilikom dužeg govora ovo stvara ometanje i otežava otvaranje usta prilikom izgovora određenih slova. Dodatno, ova maska se pomera prilikom govora. Sličan uticaj ostvaruje na artikulaciju ima i hirurška maska. Ova dva tipa maske prilikom dužeg govora stvaraju osećaj otežanog disanja, pa i to utiče na način izgovora određenih fonema. Za razliku od toga maska N95 ima čvrstu formu i ne pomera se zajedno sa ustima. Ona je odvojena fizički od usana i ne opterećuje prilikom dužeg govora. Zbog toga je i logatomska razumljivost za ovu masku najveća.

Kako bi se sagledalo da li su neki govornici usled nedovoljne koncentrisanosti ili neobučenosti za slušanje logatoma uticali na globalne rezultate izvršena je detaljnija analiza. Za svakog slušaoca izvršena je normalizacija vrednosti razumljivosti za svih 5 grupa logatoma. Normalizacija je urađena tako da su vrednosti logatomske razumljivosti za svaku od grupa podeljene sa maksimalnom vrednošću razumljivosti. Nakon toga izvršeno je usrednjavanje razumljivosti za svaku od grupa za sve slušaoce. Dobijene su sledeće vrednosti: bez maske 98.1%, maska N95 96.2%, hirurška maska 93.1%, vizir sa hirurškom maskom 84.6% i pamučna maska 80.76%. Međusobni odnos normalizovanih vrednosti je isti kao u slučaju bez normalizacije vrednosti. To znači da slušaoci nisu uticali na dobijene rezultate.

V. Zaključak

U ovom radu analiziran je uticaj zaštitne opreme koja se koristi u vreme pandemije virusa COVID19 na razumljivost govora u srpskom jeziku. Analizirana su 3 tipa maski i jedan zaštitni vizir. Pokazano je da se dugovremeni spektar govora na srpskom jeziku menja upotrebom zaštitnih maski. Na frekvencijama iznad 2 kHz zaštitne maske unose slabljenje od 3 dB do čak 10 dB. Zaštitni vizir pored slabljenja na visokim frekvencijama ističe oblast frekvencija između 400 Hz i 1000 Hz. Rezultati subjektivnih testova realizovanih slušanjem preko slušalica pokazali su da uticaj zaštitnih maski postoji. Zbog idealnih uslova slušanja dobijeni rezultati ne pokazuju uticaj zaštitne opreme u svakodnevnim uslovima. Zbog toga je sproveden subjektivni test gde su slušaoci u prostoriji slušali pripremljene grupe logatoma, koji su prethodno snimljeni sa zaštitnim sredstvima u gluvoj sobi. Rezultati su pokazali da je razumljivost najveća za masku N95, zatim slede hirurška maska, zaštitni vizir i pamučna maska. Dobijeni rezultati poklapaju se sa rezultatima iz literature dobijenim za engleski jezik. Rezultati se ne poklapaju sa vrednostima slabljenja dugovremenih spektara govora, što ukazuje da maske ne utiču samo kao filtar za govorni signal već utiču i na artikulaciju. Dobijeni rezultati pokazuju značajno smanjenje razumljivosti govornika sa određenim tipovima zaštitnih maski u reverberantnim prostorima sa ambijentalnom bukom.

ZAHVALNICA

Ovaj rad je realizovan u okviru projekta TR 36026 koga finansira Ministarstvo prosvete, nauke i tehnološkog razvoja Republike Srbije. Autori žele da se zahvale studentkinjama Maji Đaković i Melaniji Milenković na pomoći prilikom realizacije eksperimenata, kao i svima koji su učestvovali u snimanju govora i u subjektivnim testovima.

LITERATURA

[1] S. T. Jovičić, “Govorna komunikacija – fiziologija, psihoakustika i percepcija,”, pp. *29*–72, NAUKA, Beograd, Srbija, 1999.

[2] P. Bottalico, S. Murgia, G. E. Puglisi, A. Astolfi, K. I. Kirk, “Effect of masks on speech intelligibility in auralized classrooms,” *The Journal of the Acoustical Society of America*, vol. *148*, pp. *2878- 2884*, November, 2020.

[3] Y. Choi, “Acoustical measurements of masks and the effects on the speech intelligibility in university classrooms,” *Applied Acoustics*, vol. *180*, pp. *1-8*, May, 2021.

[4] N. Prodi, C. Visentin, A. Feletti, “On the perception of speech in primary school classrooms: Ranking of noise interference and of age influence,” *The Journal of the Acoustical Society of America*, vol. *133*, pp. *255- 268*, November, 2012.

[5] V. V. Ribeiro, A.P. Dassie-Leite, *et al*, “Effect of Wearing a Face Mask on Vocal Self-Perception during a Pandemic,” *Journal of Voice*, pp. *1-7*, October, 2020.

[6] M. Caniato, A. Marzi, A. Gasparella, “How much COVID-19 face protections influence speech intelligibility in classrooms?” *Applied Acoustics*, vol. *178*, pp. *1-14*, March, 2021.

[7] M. Nobrega, R. O. Mariana, M. L. Christiane, A. Nobrega, “How face masks can affect school performance,” *International Journal of Pediatric Otorhinolaryngology*, vol. *138*, pp. *1-2*, May, 2021.

[8] C. Pörschmann, T. Lübeck, J. M. Arend, “Impact of face masks on voice radiation,” *The Journal of the Acoustical Society of America*, vol. *148*, pp. *3663- 3670*, December, 2020.

[9] R. M. Corey, U. Jones, A. C. Singer, “Acoustic effects of medical, cloth, and transparent face masks on speech signals,” *The Journal of the Acoustical Society of America*, vol. *148*, pp. *2371- 2375*, October, 2020.

[10] S. R. Atcherson, B. R. McDowell, M. P. Howard, “Acoustic effects of non-transparent and transparent face coverings,” *The Journal of the Acoustical Society of America*, vol. *149*, pp. *2249- 2254*, April, 2021.

[11] S. R. Atcherson, L. L. Mendel, W. J. Baltimore, C. Patro, S. Lee, M. Pousson, M. J. Spann, “ The Effect of Conventional and Transparent Surgical Masks on Speech Understanding in Individuals with and without Hearing Loss,” *Journal of the American Academy of Audiology*, vol. *28*, no. *1*, pp. *58- 67*, 2017.

[12] M. Magee, C. Lewis, G. Noffs *et al*., “Effects of face masks on acoustic analysis and speech perception: Implications for peri-pandemic protocols,” *The Journal of the Acoustical Society of America*, vol. *148*, pp. *3562- 3568*, December, 2020.

[13] M. Vojnović, M. Mijic, “The influence of the oxygen mask on longtime spectra of continuous speech,” *The Journal of the Acoustical Society of America*, vol. *102*, no. *4*, pp. *2456-2458*, October, 1997.

[14] E. Zwicker, H. Fastl, “Psychoacoustics: Facts and Models,” 2nd edition, Springer-Verlag, Beograd, New York, 1999,

[15] *Specification for Octave, Half-Octave, and Third Octave Band Filter Sets*, ANSI S1.11, 49 CFR 227, june 2009.

[16] Tehnička dokumentacija proizvođača, dostupno na mreži: http://www.nti-audio.com/Portals/0/data/en/MiniSPL-Measurement-Microphone-Product-Data.pdf, pristupano 19.6.2021.

[17] Tehnička dokumentacija proizvođača, dostupno na mreži: http://download.steinberg.net/downloads_hardware/UR22/UR22_documentation/UR22_OperationManual_en.pdf, pristupano 19.6.2021.

[18] Tehnička dokumentacija proizvođača, dostupno na mreži: https://www.akg.com/Headphones/Professional%20Headphones/K92.html, pristupano 19.6.2021.

ABSTRACT

This paper presents an analysis of the impact of face masks on speech intelligibility. Three different types of protective masks used in pandemic conditions (cotton mask, surgical mask and N95 mask) were analyzed, as well as a protective transparent visor. Based on the speech of several people, the shape of the spectrum of long-term speech in the Serbian language in the case of the use of protective equipment was determined. Several subjective tests were performed in which the logatom intelligibility of speech with and without protective masks was measured. The experiments were organized in controlled conditions (listening through headphones) and in a room with a large reverberation time. It has been shown that the best intelligibility of speech in the Serbian language is achieved when using the N95 protective mask in relation to the other two types of analyzed masks.

Impact of COVID19 face masks on speech intelligibility in the Serbian language

Miloš Bjelić, Tatjana Miljković, Miomir Mijić, Dragana Šumarac Pavlović

Izdvajanje režima praznog hoda motora sa unutrašnjim sagorevanjem na osnovu audio zapisa

Marko Milivojčević, Emilija Kisić, Dejan Ćirić

***Apstrakt*—U ovom radu je predstavljen postupak za prepoznavanje i izdvajanje režima praznog hoda motora sa unutrašnjim sagorevanjem putničkih vozila na osnovu audio zapisa prikupljenih na ulazu u podzemnu garažu. Analiza audio zapisa je realizovana u vremenskom domenu, kako bi se omogućila obrada signala sa što manjim zahtevima u pogledu potrebne procesorske snage akvizicionog sistema sa baterijskim napajanjem. Celokupan postupak prepoznavanja i izdvajanja audio zapisa koji odgovara režimu praznog hoda motora sa unutrašnjim sagorevanjem putničkih vozila realizovan je upotrebom programskog jezika *Python*. Cilj predstavljenog postupka je priprema velikog broja prikupljenih audio signala za dalju analizu i obradu u pogledu izdvajanja karakterističnih audio obeležja i upotrebu u obuci neuronskih mreža.**

***Ključne reči*—akustičke karakteristike, audio zapis, vremenski domen, motori sa unutrašnjim sagorevanjem, akvizicija, *python*.**

I. Uvod

Primene algoritama veštačke inteligencije na audio signale vremenom postaju sve brojnije. Klasifikacija zvukova, detekcija audio događaja i prepoznavanje audio scena su primeri zadataka koji se uspešno realizuju u praksi primenom mašinskog učenja. U tom kontekstu, mašinsko i duboko učenje se mogu koristiti za prepoznavanje vrste motora sa unutrašnjim sagorevanjem na osnovu zvuka koji motori generišu. Naime, zvuk ovih motora se razlikuje u zavisnosti od pogonskog goriva koje koriste motori (benzin ili dizel). Činjenica da ljudsko uvo na osnovu zvuka koji motor generiše može da prepozna o kojoj vrsti motora se radi, poslužila je kao osnovna ideja da se napravi sistem za automatsko prepoznavanje vrste motora na osnovu zvuka koji motori generišu.

U cilju realizacije jednog ovakvog sistema najpre je bilo potrebno napraviti akvizicioni sistem za prikupljanje audio signala koji će činiti bazu na osnovu koje će se sistem obučiti nekim od algoritama mašinskog/dubokog učenja. Kako bi se formirala kvalitetna baza audio signala, bilo je neophodno prikupiti veliki broj audio signala generisanih radom motora sa unutrašnjim sagorevanjem putničkih vozila. Pomenuta banka uzoraka bi se koristila za dalju analizu i izdvajanje karakterističnih audio obeležja sa ciljem da se isključivo na osnovu akustičkih karakteristika odredi pogonsko gorivo. Pomenuti akvizicioni sistem koristi mikrofon pozicioniran u oblasti ispod motornog prostora putničkog vozila [1, 2] nakon što se detektuje prisustvo vozila.

U prethodnim fazama ovog istraživanja su analizirani spektri signala audio zapisa koji potiču od motora pokretanih različitim pogonskim gorivima pri čemu je uzet u obzir položaj mikrofona iznad poklopca motora i u oblasti ispod motornog prostora [1]. Analiziran je i uticaj položaja mikrofona u oblasti ispod motornog prostora na karakteristiku audio zapisa, gde je utvrđeno da vremenski oblik signala minimalno varira bez obzira na tačku u kojoj je postavljen merni mikrofon [2]. Ovo je omogućilo prikupljanje uzoraka bez obzira na položaj vozila u trenutku kada je ono detektovano i kada se zaustavi iznad mikrofona.

Osnovni zadatak akvizicionog sistema je prikupljanje audio zapisa motora sa unutrašnjim sagorevanjem koji se nalazi u režimu praznog hoda tj. u neopterećenom režimu, što je detaljnije opisano u [3]. I pored velikog broja uspešno prikupljenih audio snimaka, u formiranoj bazi su se pojavili snimci sa događajima koji ne sadrže samo željeni režim rada motora već i druge modove rada motora. Zbog toga je bilo potrebno da se u okviru postojećih snimaka detektuje, a zatim i izdvoji režim praznog hoda motora sa unutrašnjim sagorevanjem, što predstavlja ključni motiv ovog rada. Kako je akvizicioni sistem baziran na baterijskom napajanju, radi svoje mobilnosti, cilj je bio napraviti što jednostavniji postupak koji bi zahtevao minimalnu procesorsku snagu. Na osnovu toga je doneta odluka da se postupak izdvajanja režima praznog hoda motora bazira na obradi audio signala u vremenskom domenu, odnosno na anvelopi signala.

Rad je podeljen u više poglavlja, gde su karakteristični slučajevi prikupljenih audio signala i metod njihove obrade prikazani u poglavlju gde je opisana metodologija. U poglavlju gde su opisani rezultati data je ilustracija uspešne detekcije promene režima rada motora samo na osnovu vremenskog oblika signala, kao i nove, izdvojene signale koji predstavljaju željeni režim rada motora. Zaključci istraživanja su sumirani u poglavlju koje sledi, iza kojeg je dat spisak korišćene literature.

II. Pregled literature

Analiza dostupnih radova iz oblasti akustike koji za izvor zvuka imaju motore sa unutrašnjim sagorevanjem dala je kao rezultat podelu radova na nekoliko kategorija. Tako su

Marko Milivojčević – Akademija tehničko-umetničkih strukovnih studija Beograd, odsek: Visoka škola elektrotehnike i računarstva strukovnih studija, Vojvode Stepe 283, 11000 Beograd, Srbija (e-mail: markom@viser.edu.rs).

Emilija Kisić – Akademija tehničko-umetničkih strukovnih studija Beograd, odsek: Visoka škola elektrotehnike i računarstva strukovnih studija, Vojvode Stepe 283, 11000 Beograd, Srbija (e-mail: emilijai@viser.edu.rs).

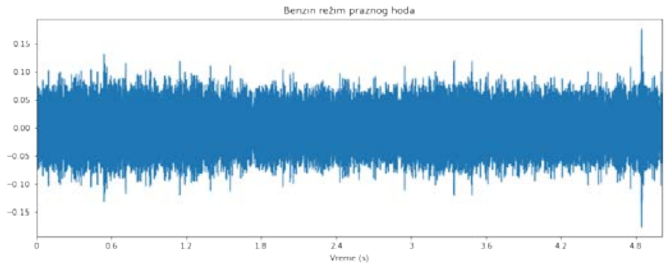
Dejan Ćirić – Univerzitet u Nišu, Elektronski fakultet u Nišu, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: dejan.ciric@elfak.ni.ac.rs).

izdvojeni radovi koji za cilj imaju detekciju-monitoring buke i brojanje vozila u saobraćaju [4, 5, 6], zatim detekciju ili predikciju nepravilnosti rada motora vozila [7, 8, 9] i klasifikaciju vozila po kategorijama (putnička, teretna, motocikli) na osnovu zvuka [10, 11]. U većini ovih radova su razvijeni namenski akvizicioni sistemi sa određenim stepenom obrade prikupljenih audio uzoraka. Osnovna razlika takvih sistema i akvizicionog sistema koji je korišćen za prikupljanje uzoraka za ovaj rad je prvenstveno pozicija i kvalitet korišćenog mikrofona, kao i nivo obrade pojedinačnih prikupljenih uzoraka pre skladištenja na memorijski medij. Posmatrajući poziciju mikrofona moguće je uočiti da se u radovima čiji je predmet detekcija-monitoring buke kao i klasifikacija vozila u saobraćaju mikrofoni pozicioniraju pored saobraćajnice, dok se kod radova fokusiranih na detekciju nepravilnosti rada motora mikrofoni pozicioniraju u oblasti iznad motornog prostora, pri čemu je poklopac motora otvoren.

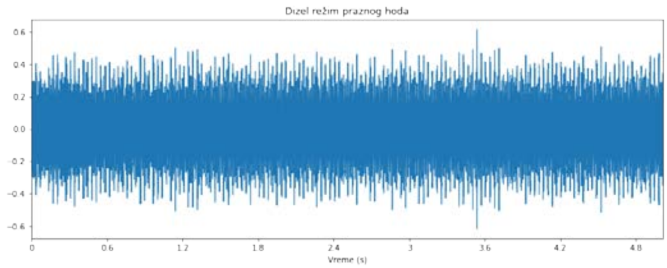
III. Metodologija

Za prikupljanje velikog broja audio zapisa motora sa unutrašnjim sagorevanjem razvijen je akvizicioni sistem koji funkcioniše autonomno. Sistem je u potpunosti mobilan jer je napajanje baterijsko, a prikupljeni zapisi se skladište na memorijsku karticu. Detekcija prisustva vozila iznad mikrofona je bazirana na ultrazvučnim senzorima koji mere rastojanje po vertikalnoj i horizontalnoj osi u različitim ravnima što onemogućava detekciju malih objekata tj. objekata koji nisu vozila. Dodatna provera, nakon detekcije ultrazvučnim senzorima, se realizuje uključenjem mikrofona i merenjem nivoa zvučnog polja. U sistemu je kreirana logika koja započinje snimanje audio zapisa tek ukoliko je detektovani zvuk iznad zadatog praga. Kako je mikrofon pozicioniran na podlozi iznad koje se vozilo krakotrajno zaustavlja dok vozač ne preuzme karticu za ulazak u podzemnu garažu i dok se ne otvori rampa, ekperimentalno je određen minimalni nivo zvučnog polja tj. prag. U sistemu je takođe definisano trajanje snimanja od 5 sekundi koje je određeno na osnovu empirijske procene zadržavanja vozila u režimu praznog hoda.

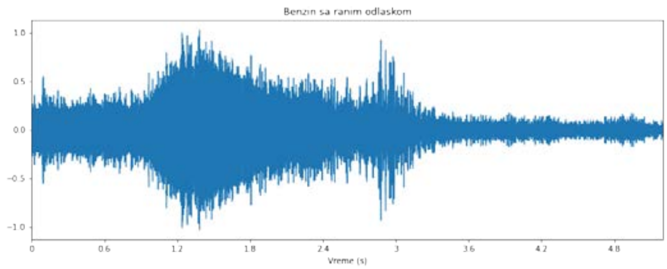
Inicijalno postavljanje sistema na ulaznu rampu podzemne garaže je pokazalo da sistem detektuje isključivo vozila i da se na audio zapisima nalaze samo signali koji potiču od motora sa unutrašnjim sagorevanjem, ali da vremena zadržavanja vozila iznad mikrofona znatno variraju. Za najveći broj vozila su dobijena dva audio snimka po vozilu pri čemu prvi snimak, bez izuzetka, predstavlja režim praznog hoda motora-stacionaran režim (Sl. 1 i 2), a drugi, odnosno u pojedinim slučajevima poslednji, snimak delimično sadrži režim praznog hoda motora nakon čega sledi povećavanje broja obrtaja radilice i režim delimičnog opterećenja motora u cilju ubrzavanja vozila (Sl. 3, 5 i 6). Na uzorku od 50 vozila ostali slučajevi nisu zabeleženi bez obzira na uticaj autoperionice koja se nalazila u neposrednoj blizini u nedeljenom zatvorenom prostoru. Dodatno je, paralelno sa autonomnim radom sistema, manuelno vođena evidencija tipa pogonskog goriva kako bi se uočila eventualna mogućnost greške tj. beleženja audio snimaka koji bi bili neupotrebljivi u odnosu na buku okoline koja je prisutna u zatvorenom prostoru i koja najvećim delom potiče od ventilacije garaže. Ovakav slučaj se nije dogodio u praksi zbog pravilno podešnog praga koji uslovljava početak beleženja snimka. Posmatrajući vremenski oblik signala, na Sl. 3, 4, 5 i 6 može se uočiti snimljena buka tek nakon odlaska vozila.



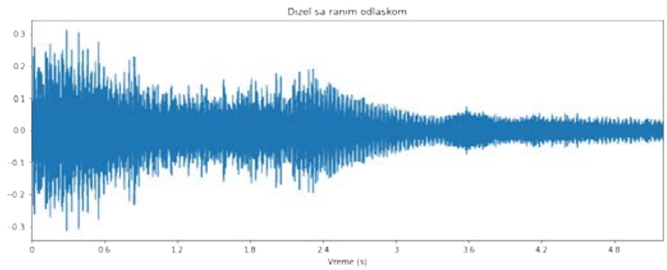
Sl. 1. Audio signal motora koji koristi benzin u praznom hodu, bez promene režima.



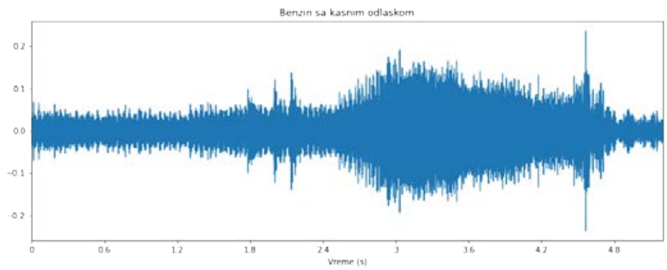
Sl. 2. Audio signal motora koji koristi dizel u praznom hodu, bez promene režima.



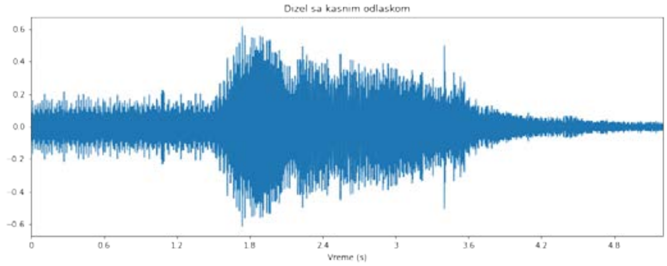
Sl. 3. Audio signal motora koji koristi benzin sa ranom promenom režima rada.



Sl. 4. Audio signal motora koji koristi dizel sa ranom promenom režima rada.



Sl. 5. Audio signal motora koji koristi benzin sa kasnom promenom režima rada.



Sl. 6. Audio signal motora koji koristi dizel sa kasnom promenom režima rada.

U cilju formiranja banke uzoraka audio snimaka stacionarnog režima rada motora i na osnovu analize vremenskog oblika signala kada vozilo odlazi, uočeno je da je moguće izdvojiti željeni režim rada „isecanjem“ dela snimka pre nego što vozilo krene da napušta poziciju tj. pređe u režim delimičnog opterećenja.

Uzimajući u obzir da se sistem napaja baterijski, da obradu celokupnog audio zapisa od 5 s radi u realnom vremenu i da se iz vremenskog oblika signala jasno mogu uočiti događaji, izdvajanje željenog režima rada je bazirano na vremenskom domenu radi smanjenja potrebne procesorske snage. Obrada u vremenskom domenu je podeljena u tri faze: formiranje anvelope amplitude signala, određivanje trenutka sečenja i isecanje signala do prethodno određenog trenutka.

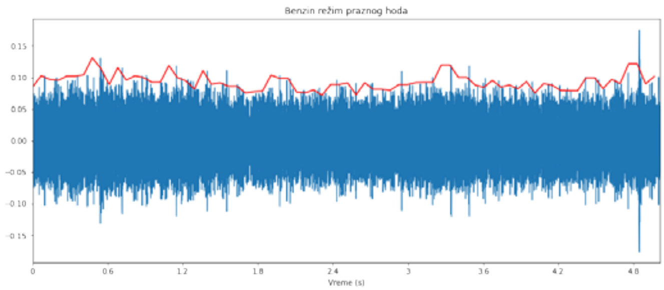
Zvučni zapis motora je sniman u audio kvalitetu 48000 odbiraka u sekundi, 16 bita po odbirku, mono. Najbolje rezultate prilikom formiranja anvelope su dali parametri sa veličinom prozora gde se beleži maksimalna vrednost signala dužine 4000 odbiraka, a pomeraj 3000 odbiraka, tj sa preklapanjem susednih prozora od 1000 odbiraka. Dobijene anvelope su prikazane na Sl. 7 do 12.

Kako bi se izdvojio samo stacionarni deo koji odgovara praznom hodu motora iz svakog snimljenog audio signala iz baze, bilo je potrebno odrediti prag (vremenski trenutak) nakon kojeg se nestacionaran deo signala odbacuje. S obzirom na prirodu problema, stacionaran deo signala se uvek javlja na početku signala (Sl. 9 do 12), tako da je bilo jasno da je prag potrebno naći u nekom trenutku nakon početka signala, odnosno u prvom trenutku kada signal prestaje da bude stacionaran. Kao što je gore već objašnjeno, slučajevi gde se prazan hod javlja kasnije (u sredini ili na kraju signala) ne postoje, tako da je bilo jasno da je potrebno odrediti samo jedan prag.

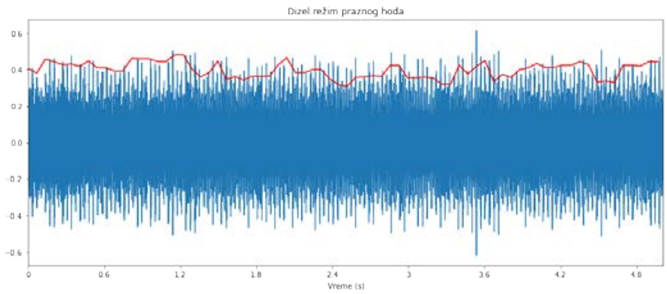
Na osnovu analize vremenskih oblika signala koji su snimljeni, došlo se do zaključka da u trenutku kada signal prestaje da bude stacionaran dolazi do naglog porasta njegove amplitude, a samim tim u tom trenutku dolazi i do veoma primetnog porasta njegove anvelope. Osnovna ideja za pronalaženje praga pomoću kojeg bi se izdvojio samo stacionaran deo signala jeste da se najpre odredi anvelopa signala kako je gore opisano, a zatim da se računa razlika između trenutne i prethodne vrednosti anvelope za sve vrednosti anvelope. Dok je signal stacionaran, očekivano je da razlika između trenutne i prethodne vrednosti anvelope bude mala. U trenutku kada signal prestaje da bude stacionaran, dolazi do skoka u anvelopi, što znači da razlika između trenutne vrednosti anvelope u trenutku kada je signal prestao da bude stacionaran i prethodne vrednosti anvelope dok je signal još uvek stacionaran mora biti znatno veća nego razlika između trenutne i prethodne vrednosti anvelope u trenucima dok je signal stacionaran. Prvi trenutak gledajući sa leva na desno (od početka do kraja signala) kada dođe do porasta razlike između trenutne i prethodne vrednosti anvelope bi trebalo da bude trenutak za postavljanje praga. Ukoliko signal anvelope označimo kao $Env(t)$, a vrednost praga kao t_L, prag možemo izračunati pomoću sledeće formule:

$$t_L = min\{Env(t) - Env(t-1) > 0.1\} \cdot \left(\frac{t_s}{N_f}\right) \quad (1)$$

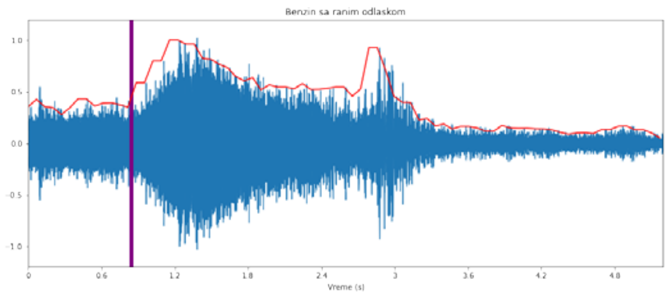
gde je sa t_s označeno trajanje celog signala, a sa N_f broj frejmova u kojima su računati maksimumi signala pomoću kojih je određena anvelopa signala. Vrednost 0.1 je određena empirijski. S obzirom na to da je potrebno postaviti prag u prvom trenutku nakon kojeg dolazi do skoka u amplitudi, gledajući od početka do kraja signala (sa leva na desno), potrebno je uzeti najmanju vrednost koja zadovoljava uslov $Env(t) - Env(t-1) > 0.1$. Takođe, kako bi se dobio tačan vremenski trenutak za postavljanje praga, a uslov $min\{Env(t) - Env(t-1) > 0.1\}$ vraća frejm anvelope u kojem dolazi do skoka u anvelopi koji označava prelazak iz stacionarnog u nestacionaran deo signala, bilo je neophodno pomnožiti dobijenu vrednost frejma anvelope sa trajanjem celog signala i podeliti sa ukupnim brojem frejmova u anvelopi. Na ovaj način dobija se željeni vremenski trenutak za postavljanje praga t_L, odnosno trenutak do kojeg se nalazi stacionaran deo signala koji želimo da izdvojimo.



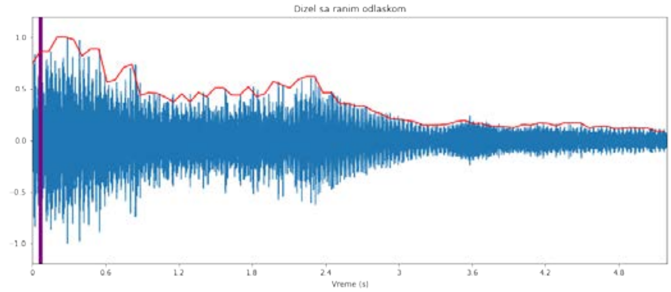
Sl. 7. Anvelopa audio signala motora koji koristi benzin u praznom hodu, bez promene režima rada.



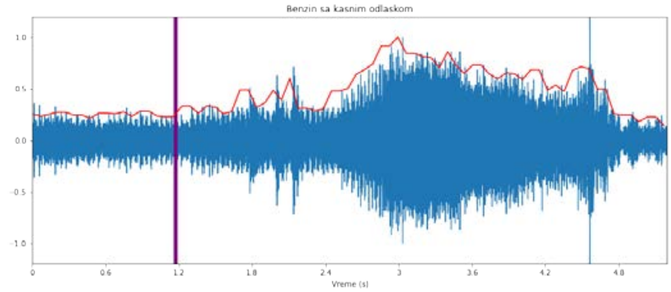
Sl. 8. Anvelopa audio signala motora koji koristi dizel u praznom hodu, bez promene režima rada.



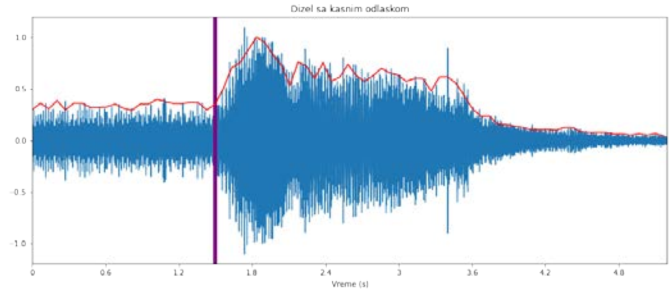
Sl. 9. Anvelopa audio signala motora koji koristi benzin sa ranom promenom režima rada (prag je označen vertikalnom linijom).



Sl. 10. Anvelopa audio signala motora koji koristi dizel sa ranom promenom režima rada (prag je označen vertikalnom linijom).



Sl. 11. Anvelopa audio signala motora koji koristi benzin sa kasnom promenom režima rada (prag je označen vertikalnom linijom).

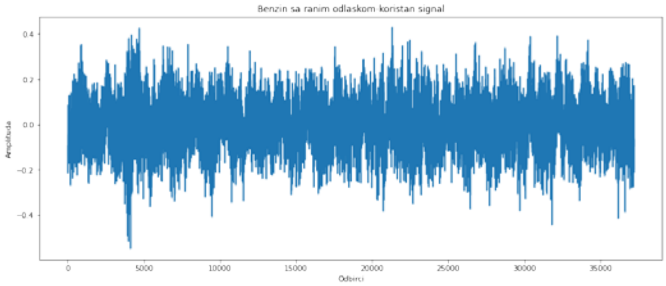


Sl. 12. Anvelopa audio signala motora koji koristi dizel sa kasnom promenom režima rada (prag je označen vertikalnom linijom).

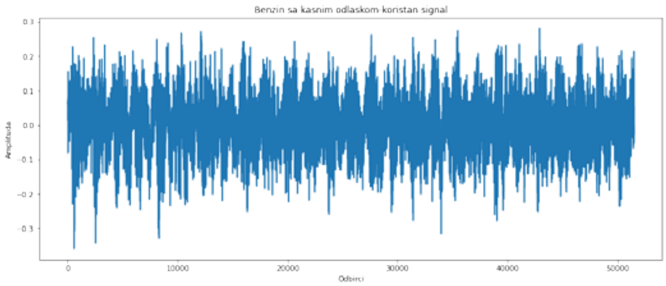
Dobijeni vremenski trenuci koji označavaju prag (trenutak do koga se izdvaja deo signala praznog hoda motora koji se koristi za banku uzoraka) je na Sl. 9 do 12 označen ljubičastom vertikalnom linijom. Na snimljenim uzorcima gde nema promene režima rada motora, vremenski trenutak odsecanja nije mogao biti određen na opisani način. U tom slučaju, čitav audio zapis se koristi kao prazan hod motora i koristi se za dalje analize i obradu.

IV. Rezultati

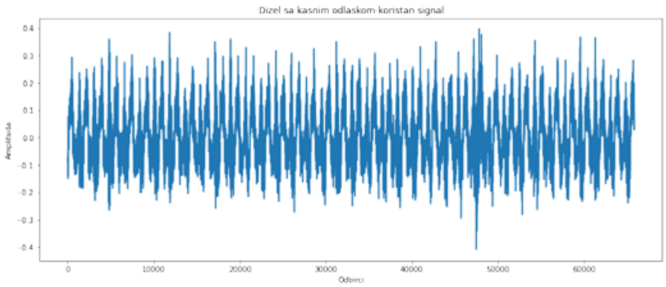
Obrada signala u vremenskom domenu, u vidu formiranja anvelope signala, kao i određivanje vremenskog trenutka kada je potrebno iseći signal na osnovu oblika anvelope, rezultirala je podatkom za svaki audio signal iz snimljene baze koji sadrži trenutak do kojeg je motor u režimu praznog hoda tj. do kojeg je signal stacionaran. Na osnovu tog podatka su iz postojećih audio zapisa upotrebom jednostavne funkcije izdvojeni željeni signali u celoj bazi. Izdvojeni signali su predstavljeni u wav formatu kao i originalni audio zapisi akvizicionog sistema. Radi ilustracije, vremenski oblici karakterističnih izdvojenih audio signala predstavljeni su na Sl. 13,14 i 15. pri čemu se na horizontalnoj osi nalazi broj odbiraka. Na Sl. 13 je prikazan izdvojeni signal koji odgovara originalnom snimku motora koji koristi benzin sa ranim odlaskom vozila, dok je na Sl. 14. takođe prikazan snimak motora koji koristi benzin samo sa kasnijim odlaskom vozila. Posmatrajući vremenski oblik signala može se uočiti da su signali veoma slični pri čemu je jedina razlika u broju odbiraka tj. trajanju izdvojenog signala, što odgovara događajima vidljivim na Sl. 9 i 11.



Sl. 13. Izdvojeni audio signal motora koji koristi benzin u praznom hodu sa ranom promenom režima rada.



Sl. 14. Izdvojeni audio signal motora koji koristi benzin u praznom hodu sa kasnom promenom režima rada.



Sl. 15. Izdvojeni audio signal motora koji koristi dizel u praznom hodu sa kasnom promenom režima rada.

Na Sl. 15. je prikazan izdvojeni signal motora koji koristi dizel pri čemu je signal izdvojen za slučaj sa kasnijim odlaskom vozila (Sl. 12.). Za slučaj kada vozilo koristi dizel i kada se dogodio rani odlazak vozila gotovo na samom početku snimka, definisana tačka sečenja je takođe bila veoma blizu početka signala (Sl. 10.) što znači da nije detektovan dovoljno dugačak stacionarni režim rada motora, pa je odgovarajućom funkcijom provere trajanja stacionarnog režima ovaj snimak odbačen bez izdvajanja signala.Opisana metoda, odnosno obrada signala u vremenskom domenu dala je jako dobre rezultate, jer je omogućila da se iz svih audio signala iz baze izdvoje željeni, korisni delovi signala, koji predstavljaju prazan hod motora, na potpuno automatizovan način, što je i bio cilj ovog rada.

V. Zaključak

Posmatrajući broj snimaka sa signalima koji predstavljaju isključivo režim praznog hoda vozila u odnosu na broj uzorkovanih vozila moguće je uvideti da je akvizicioni sistem prikupio barem jedan takav snimak za svako vozilo. Razvoj dodatne obrade zabeleženih signala, a koji ne predstavljaju isključivo režim praznog hoda motora u toku čitavog trajanja snimka, doprinelo je povećanju broja željenih audio snimaka, odnosno uvećanju banke uzoraka. Akvizicioni sistem sa dodatnom obradom signala za izdvajanje stacionarnog režima rada motora je dao identične rezultate bez obzira na tip pogonskog goriva. Upoređujući izdvojene signale na Sl. 13, 14 i 15 sa signalima koji su izvorno snimljeni kao režim praznog hoda motora (Sl. 1 i 2) mogu se videti veoma male razlike koje su rezultat vremenskog skaliranja prilikom grafičkog prikaza, odnosno jedina suštinska razlika je u dužini trajanja originalnog signala i korisnog dela signala koji je izdvojen pomoću opisane metode.

Ovakav način prikupljanja audio uzoraka, uz dodatnu obradu signala predstavljenu u ovom radu, veoma efikasno formira veliku količinu "čistih" audio snimaka. Na ovaj način napravljena je kvalitetna baza audio signala što će u budućem radu omogućiti razvoj sistema koji će nekim algoritmom mašinskog učenja, odnosno nekom od metoda klasifikacije prepoznavati o kojoj vrsti motora se radi.

ZAHVALNICA

Prezentovano istraživanje je realizovano zahvaljujući gospodinu Neđi Petijeviću ispred firme Novi Dom doo u Beogradu koji je omogućio pristup ulaznoj rampi podzemne garaže uz poštovanje svih bezbednosnih procedura. Ovaj rad je podržan od strane Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije.

LITERATURA

[1] M. Milivojčević, F. Pantelić, D. Ćirić, "Comparison of frequency characheristic of sound generated by internal combustion engines depending on fuel," Proc. 26th Noise and Vibration, Niš, Serbia, pp. 115-120, 6-7 December 2018.

[2] M. Milivojčević, F. Pantelić, D. Ćirić, "Pozicioniranje mikrofona prilikom snimanja audio karakteristika motora putničkih vozila" Proc. 63rd National Conference on Electrical, Electronic and Computing Engineering ETRAN, Srebrno Jezero, Serbia, pp. 58-62, 3-6 June 2019.

[3] M. Milivojčević, D Ćirić, „Izdvajanje značajnih akustičkih karakteristika motora sa unutrašnjim sagorevanjem"Proc. 64rd National Conference on Electrical, Electronic and Computing Engineering ETRAN, Belgrade, Serbia, 2020.

[4] J. F. Forren and D. Jaarsma, "Traffic monitoring by tire noise," Proc. Conference on Intelligent Transportation Systems, pp. 177-182, 1997.

[5] R. Lopez-Valcarce, C. Mosquera, R. Perez-Gonzalez, "Estimation of road vehicle speed using two omnidirectional microphones: A maximum likleihood approach", *EURASIP J. Appl. Signal Process.*, p.1059, 2004.

[6] B. Barbagli, G. Manes, R. Facchini, "Acoustic sensor network for vehicle traffic monitoring," Proc. 1st International Conference on Advances in Vehicular Systems, Technologies and Applications, Venice, Italy, pp. 1–6, 24–29 June 2012.

[7] V. Singh, N. Meena, "Engine fault diagnosis using DTW, MFCC and FFT," Proc. First International Conference on Intelligent Human Computer Interaction (IHCI 2009), Allahabad, India, pp. 83-94, 20-23 January 2009.

[8] S. K. Lee, S. J. Kim, "Internal combustion engine sound-based fault detection and diagnosis using adaptive line enhancers." *Proc. Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 222, no. 4, pp. 593–605, Apr. 2008.

[9] J.L. Firmino, J.M. Neto, A.G. Oliveira, J.C. Silva, K.V. Michina, M.C. Rodrigues, "Misfire detection of an internal combustion engine based on vibration and acoustic analysis," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, no. 43, art. no. 336, 2021.

[10] J. George, L. Mary, Riyas K S, "Vehicle detection and classification from acoustic signal using ANN and KNN," Proc. Int. Conf. on Control Communication and Computing (ICCC), pp. 436-439, 2013.

[11] Y. Na, Y. Guo, Q. Fu, Y. Yan, "An acoustic traffic monitoring system: design and implementation," 2015 IEEE 12th Int. Conf. on Ubiquitous Intelligence and Computing and 2015 IEEE 12th Int. Conf. on Autonomic and Trusted Computing and 2015 IEEE 15th Int. Conf. on Scalable Computing and Communications and its Associated Workshops, pp. 119-126, 2015.

ABSTRACT

In this paper, procedure for recognition and extraction of idling mode of the internal combustion engines of passenger vehicles based on audio recordings acquired at the entrance of underground parking is presented. Audio recording analysis is carried out in the time domain in order to have signal processing with as little as possible requirements regarding necessary processing power of the acquisition system with battery power supply. Overall procedure for recognition and extraction of useful parts of audio recordings related only to idling mode of the internal combustion engines of passenger vehicles is realized using *Python* programming language. The main goal of the presented procedure is the preparation of large number of acquired audio signals for further analysis and processing in the context of extraction of characteristic audio features and application of these features in neural network training.

Extraction of the idling mode of the internal combustion engine based on the audio recording

Marko Milivojčević, Emilija Kisić, Dejan Ćirić

Whispered Speech Recognition Based on DTW algorithm and μFCC feature

Branko R. Marković and Jovan Galić

Abstract—**This paper presents the results of normal and whispered speech recognition using the μFCC (μ-law Frequency Cepstral Coefficients) feature. This feature uses a warping frequency function and it is applied at the front-end of ASR. The Dynamic Time Warping algorithm is used at the back-end of the ASR system. All experiments were performed using the part of the Whi-Spe database. Four scenarios are analyzed: normal/normal, whisper/whisper, normal/whisper and whisper/normal in the speaker dependent mode. The results confirmed an expected improvement in recognition of whispered speech compared to the standard LFCC and MFCC features.**

Keywords— μFCC (μ-law Frequency Cepstral Coefficients); whispered speech; DTW (Dynamic Time Warping); speech recognition.

I. Introduction

THE speech has different modes and one of standard classification is: whisper, soft speech, normal (neutral), loud and shout [1]. Very interesting is whispered speech because is quietly different compared to normal, and at same time is intelligible and in most cases easy to understand. A lot of researches who were involved in normal speech recognition also are trying to apply different tools for whisper [2-4]. They made different results with more or less success.

This paper analyzes different scenarios related to whispered and normal speech. For this purpose the DTW algorithm is used with specific warping scale (μ warping**).**

The DTW (Dynamic Time Warping) algorithm [5] is known as "old" pattern matching method for back-end ASR systems. There are many different new method like HMM (Hidden Markov Models), DNN (Deep Neural Networks), SVM (Support Vector Machines) etc. but for quick and valuable compression DTW is still very successful. Many researchers use the DTW as a method for initial classification of patterns, and then use other methods for more precise results.

For this research speech patterns from the Whi-Spe database [6] are used. The database contains 10,000 patterns which are representation of 50 different words spoken in normal and whispered mode. Five male and five female volunteers were included in this recording. For recording the special acoustical room is used where noise is suppressed.

All experiments are based on three types of features: Mel-frequency cepstral coefficients (MFCC) plus delta, Linear frequency cepstral coefficients plus delta and μFCC cepstral coefficients (LFCC) plus delta. For all experiments the following training/test scenarios are examined: comparison between normal and normal patterns (N/N scenario), comparison between whisper and whisper patterns (W/W), comparison between normal and whisper patterns (N/W) and comparison between whisper and normal patterns (W/N).

The paper has the following structure: the second part explains how to obtain MFCC and LFCC feature vectors from the initial wave files. The third part explains how to obtain μFCC feature vectors. The forth part shows the results of experiments for all mentioned features. The final remarks and hints for further research are presented at the conclusion.

II. MFCC and LFCC Features Extraction

Mel-frequency cepstral coefficients are traditionally very popular feature for speech characterization. The mel-frequency scale (Fig. 1) emulates human's ear perception. The frequency in mel is calculated using the following equation:

$$f[mel] = 2595 * \log_{10}(1 + f[Hz]/700) \quad (1)$$

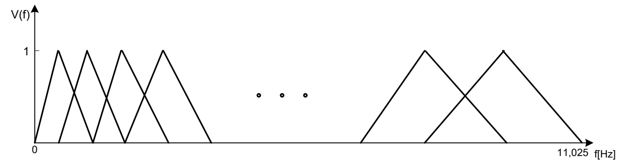


Fig. 1. Filters based on mel scale

Linear-frequency scale has the same shape for all filters (Fig. 2). The Linear frequency cepstral coefficients shows some advantage compare to MFCC in case of speaker identification in whisper [7].

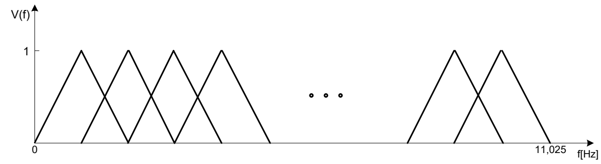


Fig. 2. Filters based on linear scale

The way to obtain MFCC and LFCC features is depicted in Fig. 3. It is a process of getting usually three types of vectors as an output: vectors of cepstral coefficients, vector of cepstral and delta cepstral coefficients and vectors of cepstral, delta

Branko R. Marković is with the Faculty of Technical Science Čačak, University of Kragujevac, Department of Computer Science and Software Engineering, Čačak, Svetog Save 65, Serbia (e-mail: brankomarko@ yahoo.com)

Jovan Galić is with the Faculty of Electrical Engineering, University of Banja Luka, Department of Electronics and Telecommunications, Banja Luka, Bosnia and Herzegovina (e-mail: jovan.galic@etf.unibl.org)

cepstral and delta-delta cepstral coefficients. For this research the first two types are used.

The inputs are wave files from Whi-Spe database [6]. All patterns are recorded with sampling rate of 22050 Hz, 16 bits per sample. Only difference between MFCC and LFCC features is in the scale which is used: MFCC counts the log energy over mel scale while LFCC counts the log energy over linear scale.

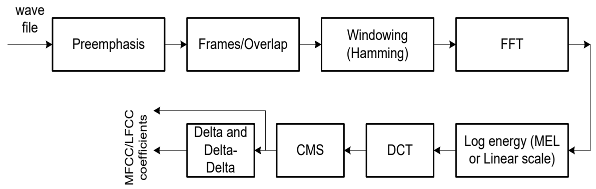


Fig. 3. Block diagram for MFCC/LFCC based features

The preprocessing assumes a several steps to get the feature vectors from an initial wave file (Fig. 3). The following steps are performed: preemphasis, framing with overlap, Hamming window, Fast Fourier Transformation (FFT), Log energy over a specific scale, DCT (Discrete Cosine Transform) and CMS (Cepstral Mean Subtraction) normalization.

The first is preemphasis block and it produces a spectrally flattened signal. Then, in the framing/overlap block, the signal as an output of preemphasis, is divided into frames. Each frame contains 512 samples, and then it is overlapped 50%. In next block the frames are weighted with the Hamming window.

The next step is the FFT, which calculates short time spectra of the signal. Then, the Log energy is calculated over the specific scale (mel or linear).

Finally, the Discrete Cosine Transformation with CMS are applied to produce the cepstral coefficients.

The CMS is a normalization method and is very important approach for whispered speech recognition [10,11].

For calculation of the first derivative (Delta), three neighboring frames are included.

Based on the mel scale and the preprocessing, two types of vectors are produced:

- vector containing 12 MFCCs and
- vector containing 24 coefficients (12 MFCCs and 12 Delta MFCCs).

Similarly, based on the linear scale the following vectors are obtained:

- vector containing 12 LFCCs and
- vector containing 24 coefficients (12 LFCCs and 12 Delta LFCCs).

These types are used in all experiments.

III. μFCC Feature Extraction

The reason to involve μFCC feature is in the following: due to unvoiced nature of whispered speech the spectrum is relatively flat. Some significant part of whispered information is in higher part of speech spectrum. The mel scale, due to its nature, is not able to "catch" these information. The linear scale shows better performances for part of higher frequencies but has worse resolution for lower frequencies. So, as a compromise, the new, warping function of frequency is involved [8,9].

This function is called μ-law and originally is involved for speech compression and expending in Japan and North America. The μ-law is defined by the following equation:

$$f_\mu = f_N \frac{\ln(1+\mu \frac{f}{f_N})}{\ln(1+\mu)} \qquad (2)$$

where $f_N = f_S/2$ (f_S is the sampling frequency), and for these experiments f_N =11025 Hz. The μ is a positive number and can have different values. For this research μ takes values {0,1,2}. Fig. 4 shows the warping functions of μ-law for these three values {0,1,2}[9]. For μ =0, the scale is linear, and practically the feature are LFCC, as mentioned before.

The μFCC feature should make some compromise between linear and mel scale with focus to improve whispered speech recognition.

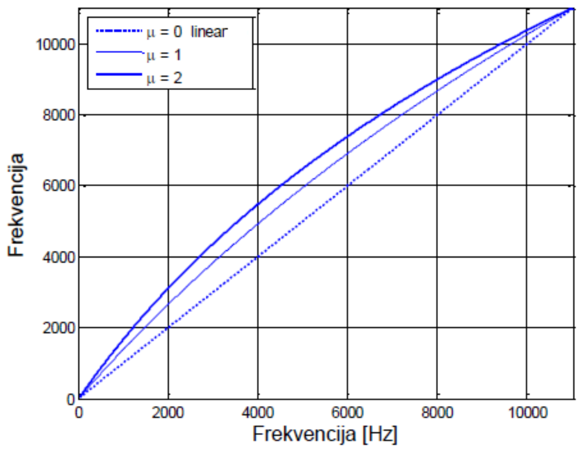


Fig. 4. Warping functions

In order to obtain μFCC features the block diagram from Fig. 5. is used. It is clone to earlier mention diagram for LFCC and MFCC feature (Fig. 3). The main difference is usage of warping function over frequencies when Log energy is calculated.

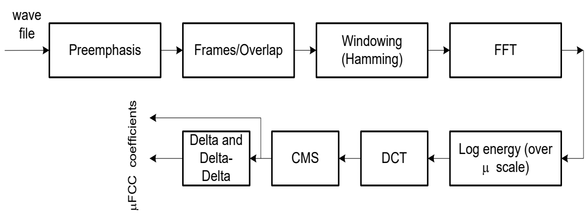


Fig. 5. Block diagram for μFCC based features

For these experiments two values for μ are used: 1 and 2. So, for μ=1 two types of vectors are considered:

- vector containing 12 μFCCs (μ-law Frequency Cepstral Coefficients) and
- vector containing 24 coefficients (12 μFCCs and 12 Delta μFCCs).

Similarly, for μ=2 two types of vectors are produced. All these vectors are used in all experiments.

IV. Results

For the purpose of testing these different features a software package is developed using the MATLAB. There are two parts of this software: the first one converts the targeted wave files from the Whi-Spe database into the set of MFCC, LFCC, and μFCC feature vectors (two type of vectors for all of them, and for μFCC two different values of μ are used). The second part compares feature vectors using the DTW algorithm.

The DTW algorithm uses the dynamic programming and it allows finding an optimal path between the starting and ending points. The speech patterns are represented by a set of feature vectors. Two set of patterns are used: the first set of 50 patterns is used as a reference, and all other patterns (nine sets, each of 50 patterns) are used as test data. For a local constraint the type I is implemented [12]. No global constraints are used.

This research used two (of ten) speakers from Whi-Spe database: one female (Speaker1) and one male (Speaker6).

For all types of feature vectors mentioned before, the results are expressed as the Word Recognition Rate (WRR). Four scenarios in speaker dependent mode are analyzed: normal/normal (denoted as N/N), whisper/whisper (W/W), normal/whisper (N/W) and whisper/normal (W/N). Tables I and II shows results for “Speaker1” and “Speaker6” using four different vector features with 12 cepstral coefficients.

TABLE I
WORD RECOGNITION RATE FOR “SPEAKER1” USING 12 CEPSTRAL COEFFICIENTS

Scenario	LFCC	μFCC (μ=1)	μFCC (μ=2)	MFCC
N/N	98.89	99.11	99.56	99.78
W/W	95.56	96.89	97.11	97.78
N/W	81.56	84.75	85.11	78.22
W/N	67.78	66.44	64.00	46.44

TABLE II
WORD RECOGNITION RATE FOR “SPEAKER6” USING 12 CEPSTRAL COEFFICIENTS

Scenario	LFCC	μFCC (μ=1)	μFCC (μ=2)	MFCC
N/N	96.44	98.22	98.67	99.33
W/W	89.33	92.89	95.11	95.11
N/W	62.00	65.56	68.00	68.00
W/N	51.33	50.89	49.78	36.44

Tables III and IV give results for “Speaker 1” and “Speaker 6” using all mentioned feature vectors with 12 cepstral and 12 delta cepstral coefficients.

TABLE III
WORD RECOGNITION RATE FOR “SPEAKER1” USING 12 CEPSTRAL AND 12 DELTA CEPSTRAL COEFFICIENTS

Scenario	LFCC	μFCC (μ=1)	μFCC (μ=2)	MFCC
N/N	98.67	98.89	99.56	99.56
W/W	96.00	97.11	97.11	97.78
N/W	82.00	84.44	84.89	78.22
W/N	66.44	68.00	64.44	45.78

TABLE IV
WORD RECOGNITION RATE FOR “SPEAKER6” USING 12 CEPSTRAL AND 12 DELTA CEPSTRAL COEFFICIENTS

Scenario	LFCC	μFCC (μ=1)	μFCC (μ=2)	MFCC
N/N	96.22	97.78	98.44	99.11
W/W	88.67	93.11	94.22	94.89
N/W	63.11	66.44	67.33	67.56
W/N	51.33	50.89	49.33	37.56

Based on the results from Tables I and II can be concluded that the MFCC feature gives very good results for match scenarios (Normal/Normal and Whisper/Whisper). But for mismatch scenarios μFCC feature is giving better results than MFCC (about 8% for N/W scenario, 30% for W/N scenario – for Speaker1 and about 28% for W/N scenario - for Speaker6). Also, for match scenarios μFCC feature gives better result than LFCC feature for both speakers.

The results in Tables III and IV are based on vectors with 12 cepstral coefficients plus 12 delta cepstral coefficients. In some cases there are improvements related to the word recognition rate with these vectors, but they are not significant (i.e. for Speaker1 and W/N scenario, μFCC feature (μ=1) gives 1,5% better result than for cepstral).

In general, based on results from Tables I-IV it is easy to conclude that the Speaker1 has better results in all scenarios compared to Speaker6. Speaker1 is better “whisperer”. Hence, on Fig. 6. the results of Speaker1 are depicted for the feature vectors which contains 12 cepstral coefficients.

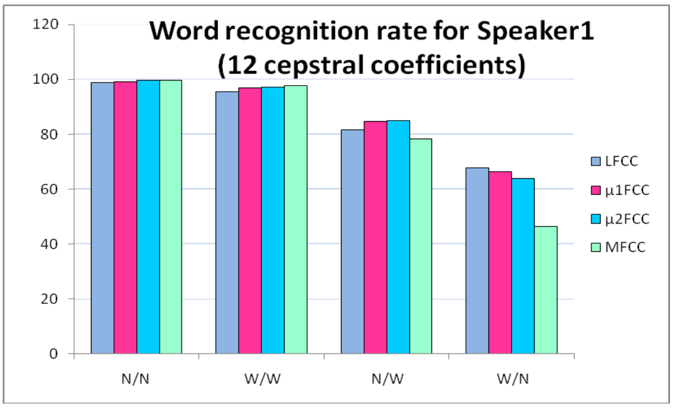


Fig. 6. WRR for Speaker1 using vectors of 12 cepstral coefficents

As it expected, the best results are for N/N, W/W, N/W and W/N scenarios, respectively. With µ1FCC the µFCC feature where µ=1, is denoted. Similarly, µ2FCC means µ=2.

It is interesting that µFCC feature for N/W scenario gives better results than LFCC and MFCC.

Fig. 7 shows results of Speaker1 for all scenarios when vectors are containing 12 cepstral plus 12 delta cepstral coefficients.

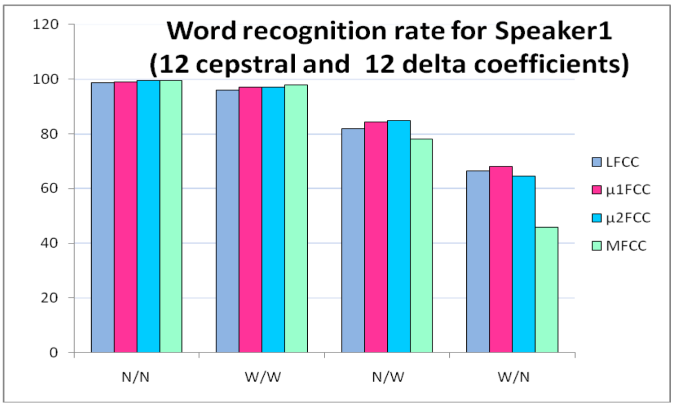


Fig. 7. WRR for Speaker1 using 12 cepstral plus 12 delta coefficents

If Fig. 6 and Fig. 7 are compared the similar trend for all scenarios and all features is evident. Only, for W/N scenario LFCC and µ1FCC changed their places.

V. Conclusion

As it expected, the best recognition results are obtained for Normal/Normal scenario and they are above 99% when MFCC feature is used. Also for Whisper/Whisper scenario the WRR is the best with MFCC. So, for match scenarios MFCC gives good results.

When mismatch scenarios are analyzed µFCC allows better results than MFCC. This is especially visible for W/N scenario where the improvement is from 28% to 30%. Obviously, these results are optimistic and give a hint to make more detailed research how the values of µ cause different WRR.

Comparing the length of feature vectors (12 cepstral coefficients vs. 24 coefficients -12 cepstral and 12 delta) the results are similar. The reason behind it can be "clean" speech in Whi-Spe database, while the delta parameters are usually efficient for noisy speech.

Further analysis may include all ten speakers from Whi-Spe database, and also more different values for µ. Instead of {0,1,2} values it can be numbers with decimal point [9]. That should provide new interesting results.

References

[1] C. Zhang, J.H.L. Hansen, "Analysis and classification of Speech Mode: Whisper through Shouted," *Interspeech 2007*, 2007, pp. 2289-2292

[2] S.T. Jovičić, Z.M. Šarić, „Acoustic analysis of consonants in whispered speech," *Journal of Voice*, 22(3), 2008, pp. 263-274.

[3] Matsuda M., Kasuya H. (1999). "Acoustic nature of the whisper", *Proc. Eurospeech 99*, 1, 1999, pp. 137-140.

[4] S.T. Jovičić, "Formant feature differences between whispered and voiced sustained vowels," *ACUSTICA - Acta Acustica*, 84(4), 1998, pp. 739-743.

[5] L. Rabiner, B-H. Juang, "Fundamentals of speech recognition", Prentice Hall, New Jersy, 1993

[6] B. Marković, S.T. Jovičić, J. Galić, Đ. Grozdić,: Whispered Speech Database: Design, Processing and Application, 16th International Conference, TSD 2013, I. Habernal and V. Matousek (Eds.): TSD 2013, LNAI 8082, Springer-Verlag Berlin Heidelberg, pp. 591-598

[7] Fan X., Hansen J.H.L. Speaker identification with whispered speech based on modified LFCC parameters and feature mapping. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2009). 2009. pp. 4553–4556.

[8] Sklar B. Digital Communications: Fundamentals and Applications: 2nd edition. Prentice-Hall. 1988. 776 p.

[9] J. N. Galić, S. T. Jovičić, V. D. Delić, B. R. Marković, D. S. Šumarac Pavlović, Đ. T. Grozdić, *„HMM-based Whisper Recognition using µ-Law Frequency Warping "*, SPIIRAS Proceedings, Issue No 3(58), 2018, pp. 27-52, ISSN 2078-9181, DOI 10.15622/sp.58

[10] J. De Veth, L. Boves, "Channel Normalization Techniques for Automatic Speech Recognition over the Telephone", – Speech Communication, 25, pp. 149-164, 1998.

[11] Grozdić Đ., Jovičić S., Šumarac-Pavlović D., Galić J., Marković B. (2017). "Comparison of Cepstral Normalization Techniques in Whispered Speech Recognition", Advances in Electrical and Computer Engineering, Vol. 17. Number 1, 2017, pp 21-26.

[12] Sakoe H. and Chiba S. (1978). „Dynamic programming optimization for spoken word recognition", IEEE Trans. Acoustics, Speech, Signal Proc., pp 43-49, 1978

The Experiments in SVM-based Whispering Speaker Identification

Jovan Galić, Branko Marković and Đorđe T. Grozdić

***Abstract*—This paper presents results of automatic speaker recognition in normally phonated (neutral) and whispered speech, based on Support Vector Machines (SVM) and Whi-Spe speech database. The performance of the recognizer is examined in matched N/N (Neutral/Neutral) and W/W (Whispered/Whispered) train/test scenarios for different types of kernels (Radial basis function, Polynomial, Linear, and Sigmoid). The best accuracy is obtained with a polynomial kernel (96,12% for neutral speech and 92,16% in case of whispering). The influence of the size of training data on the performance of the recognizer is examined, as well.**

Keywords—Speaker recognition; Whispered speech; Whi-Spe database; MFCC; SVM algorithm.

I. Introduction

The technology of automatic speech and speaker recognition has made significant progress in the last two decades. Still, some disadvantages remain. The key imperfection is the considerable degradation of the performance in adverse conditions [1]. As well, speech technologies are designed for recognition of the most commonly used mode of phonation, i.e. neutral speech. Speech mode can be classified into the five main categories: whispered speech, soft speech, normally phonated speech (neutral speech), loud speech and shouted speech [2].

Nowadays, whisper is often used in a daily life, especially over the mobile phones. There are multiple reasons to use whisper: when someone doesn't like to disturb others, when the loud speech is prohibited or unpleasant, when the information to speak is secret, when someone wishes to hide identity etc. Also, whisper can be produced due to health problems: it may happen after laryngitis or rhinitis [3]. Whisper as a speech mode is characterized by a lack of glottal vibration, noisy excitation of the vocal tract and in general, the changes of the vocal tract structure.

Jovan Galić is with the Faculty of Electrical Engineering, University of Banja Luka, Department of Electronics and Telecommunications, Banja Luka, Patre 5, Bosnia and Herzegovina (e-mail: jovan.galic@etf.unibl.org)

Branko Marković is with the Faculty of Technical Science Čačak, University of Kragujevac, Department of Computer Science and Software Engineering, Čačak, Svetog Save 65, Serbia (e-mail: brankomarko@yahoo.com).

Đorđe T. Grozdić is with Grid Dynamics Holdings, Inc, Blvd. Mihajla Pupina 115, Belgrade, Serbia; and School of Electrical Engineering, University of Belgrade, Blvd. Kralja Aleksandra 73 (e-mail: djordjegrozdic@gmail.com).

There are main differences between neutral and whispered speech. It was determined that formant frequencies for whispered vowels are substantially higher than for the neutral voice [4]. Also, compared to neutral speech, whisper has less energy, longer durations of speech and silence intervals, flatter spectrum and lower sound pressure level (SPL) [2]. Despite of these "weaknesses", the intelligibility of whisper is pretty high [5]. But, non-linguistic information (like age, sex, emotions or identity), is still a big challenge for research in whispered speech.

The oscillogram and spectrogram of the short sentence "Govor šapata." ("Whispered speech" in English), uttered in neutral and whispered speech are depicted in Figures 1 and 2, respectively. Because of the lack of sonority, the difference in amplitude intensities can be observed, especially for vowels [6]. The analysis of spectrograms shows that some parts of the spectrum are well preserved in whisper, especially in the case of unvoiced consonants and plosives. Moreover, the spectrogram shows that the harmonic structure of vowels is lost in the case of a whisper [7].

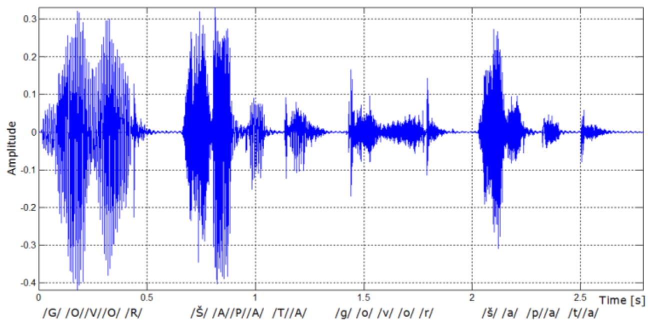


Fig. 1. The waveform of sentence "Govor šapata" in neutral phonation (capital letters) and whispered phonation (small letters).

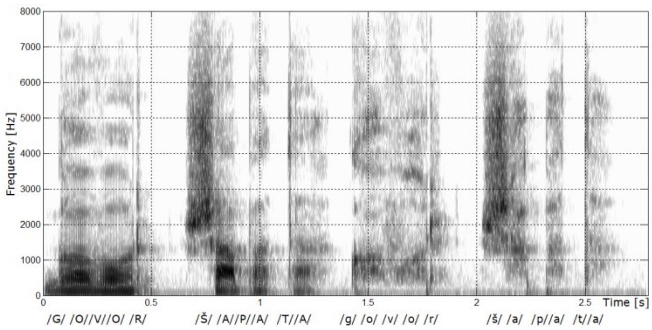


Fig. 2. The spectrogram of sentence "Govor šapata" in neutral phonation (capital letters) and whispered phonation (small letters).

Whispering speaker identification is a great challenge for state-of-the-art speaker recognition systems. In a range of speech modes from whisper to shouted, whispered speech has the most negative influence on the performance of Automatic Speech/Speaker Recognition (ASR) systems [2]. Recently, the use of Gaussian Mixture Models and K-means algorithms has been analyzed in whispering speaker ID for mel and exponential frequency scales [8]. Also, formant gap features showed higher accuracy in speaker verification compared to baseline features [9].

Automatic speaker recognition can be classified into identification and verification. Methods for speaker ID can be divided into text-independent and text-dependent [10]. For a text-independent ASR system, models for a particular speaker are irrespective of uttered speech, whereas in a text-dependent system the performance of speaker recognition depends on uttered phrases. In this research, text-dependent closed set speaker identification based on Support Vector Machines (SVM) and Mel-Frequency Cepstral Coefficients (MFCC) was analyzed.

The goal of the study presented in this paper is to analyze speaker identification accuracy for neutral and whispered speech, and classification based on SVM. This paper is organized in the following manner. In Section II the classification based on SVM is shortly discussed. The basic characteristics of the ASR system and speech database used for speaker recognition are described in Section III. The results of conducted experiments are given in Section IV whereas concluding remarks and directions for future research are stated in Section V.

II. Support Vector Machines

The SVM classifier is a relatively simple machine-learning algorithm that minimizes the structural risk [11]. Initially, the SVM classifier was introduced for linearly separable classes of objects. The separation of classes is obtained with an n-dimensional hyperplane that maximizes the margin between classes (circles and squares), as depicted in Figure 3. The margin is labeled as *M* and support vectors are hatched.

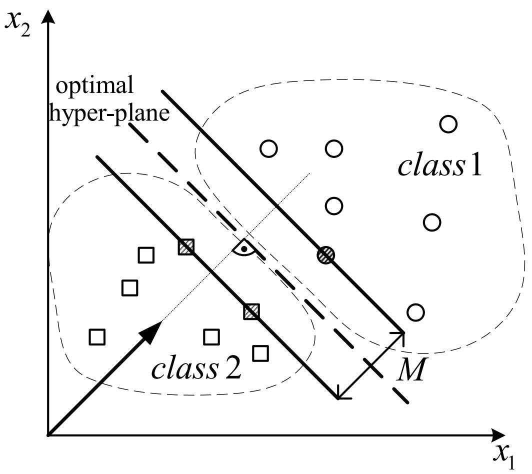


Fig. 3. Determination of hyperplane in SVM for linearly separable classes.

However, the classes are not linearly separable in most practical applications. To overcome that limitation, a non-linear transformation is performed on a feature vector. The mapping into high-dimensional feature space (in which linear separation is expected) is performed by using kernel function. Each function that satisfies necessary properties (Mercer's theorem) can be used as a kernel [12]. The most used types of kernels are:

- Radial basis function kernel (adjustable parameter γ)

$$K(x_1,x_2)=\exp(-\frac{\|x_1-x_2\|^2}{2\sigma^2});\ \gamma=\frac{1}{2\sigma^2}. \quad (1)$$

- Polynomial kernel (adjustable parameters are the slope α, the constant term c and the polynomial degree d)

$$K(x_1,x_2)=(\alpha x_1^T x_2+c)^d. \quad (2)$$

- Linear kernel (adjustable parameter c)

$$K(x_1,x_2)=x_1^T x_2+c. \quad (3)$$

- Hyperbolic Tangent (Sigmoid) kernel (adjustable parameter are α and the constant term c)

$$K(x_1,x_2)=\tanh(\alpha x_1^T x_2+c). \quad (4)$$

Because SVM is a static classifier, meaning that it works with fixed-size input data, the application in speech/speaker recognition has some restrictions. Some hybrid solutions were developed to overcome that limitation [13]. Another issue is multiclass classification, which is commonly solved by using one of the two following techniques. The first technique includes comparison of each class against all the others (one-vs-all) and the second technique confronts each class against all the others separately (one-vs-one). In this study, a one-vs-all comparison strategy is used.

III. System for Recognition

A. Speech database

One of the real problems related to the whispered speech research is a shortage of an extensive speech databases. There are some of them developed so far [14-17] for Japanese, Mandarin and English. In order to do this research, the Serbian speech database called Whi-Spe (abbreviation of *Whispered Speech*) is used [18]. The database contains two parts: the first one has recordings of whispered words, and the second one has recordings of the same words uttered with neutral speech. The vocabulary of 50 different words is divided in three groups: basic colors (6 words), numbers (14 words) and phonetically balanced (30 words). For recordings of the Whi-Spe ten volunteers (5 female and 5 male) uttered the vocabulary ten times in both speech modes, neutral and whisper. Hence, the speech database contains 10.000 represents of words in form of wave files and the total duration is 2 hours.

More details about the Whi-Spe database regarding segmentation procedure and quality control can be found in [18].

B. ASR system

Because SVM-based classifiers need feature vectors of fixed dimension, the variation in the duration of input speech utterances must be uniform. The two most common approaches for making a fixed number of frame windows for SVM classifier are using variable window size (with constant overlapping factor) and fixed window size (with variable overlapping factor). This causes some loss of information, especially in long speech utterances. In this paper, segmentation based on variable window size is chosen, using 13 overlapping windows, same as in the SVM-based speech recognition [19].

C. Feature vector extraction

The most common features used in ASR systems are Mel Frequency Cepstral Coefficients (MFCC). The diagram for obtaining the MFCC feature vector is depicted in Fig. 4.

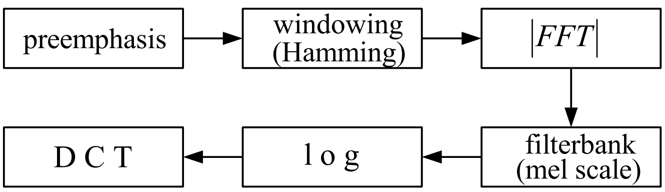


Fig. 4. The diagram for MFCC feature vector generation.

The generation of MFCC feature vectors includes the following steps: preemphasis, framing with overlap and Hamming windowing, application of the Fast Fourier Transformation (FFT), using the mel scale, calculating log energy and finally obtaining the cepstral, delta cepstral and delta-delta cepstral coefficients (based on Discrete Cosine Transformation - DCT).

The MFCC feature vector is obtained by using static features (13) along with time derivatives (delta and delta-delta) and cepstral mean normalization (39 in total). Lastly, each speech utterance from the database is represented with a vector of 507 coefficients (13 x 39) and later used as an input to the SVM classifier. The speech recognizer is developed with Python (version 3.6) using the Scikit learn package.

IV. Results

In the initial experiment baseline speaker ID performance are evaluated in 2 matched train/test scenarios: N/N and W/W. In order to have a more reliable evaluation of the performance, 10-fold cross-validation was done. The average accuracy is used as a metric for performance evaluation.

Firstly, the influence of kernel selection on recognizer performance was analyzed. The experiments were done for 4 kernels: radial basis function (RBF), polynomial (with degree *d*=3), linear and sigmoid.

The results are graphically presented in Figure 5.

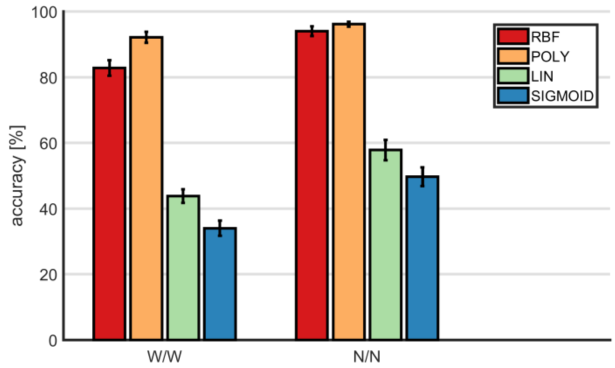


Fig. 5. Results of speaker identification (average accuracy with standard deviation) in neutral (N/N) and whispered speech (W/W) depending on type of kernel.

As it can be seen from Figure 5, the best results in whispering speaker ID are obtained for the polynomial kernel (92,16%). As well, the best result in recognition of speakers in neutral phonation is obtained using the polynomial kernel (96,12%). The experiments showed that speaker identification performance for RBF kernel is poor compared to poly kernel. Linear and sigmoid kernels are practically useless and show poor performance, especially for speaker identification in whisper mode.

As well, the influence of the percentage of the database used for training (i.e. number of training instances) on classification performance was examined. In this experiment, the polynomial kernel was used because it showed the best recognition in the previous experiment. The results are presented in Figure 6. An important part of bar graph is emphasized (higher than 75%).

For the speaker recognition in neutral mode the recognition performance starts from 94,71% (for training 75% of full capacity database, i.e. 3750 utterances) and reaches final 96,12% (for 95%, i.e. 4750 utterances). On the other hand, whispering speaker recognition performance is in the range of 89,49% (for 75%) up to 92,16% (for 95%). As observed, the saturation effect can be seen for both neutral and whisper speech modes (saturation effect is less for whispered speech).

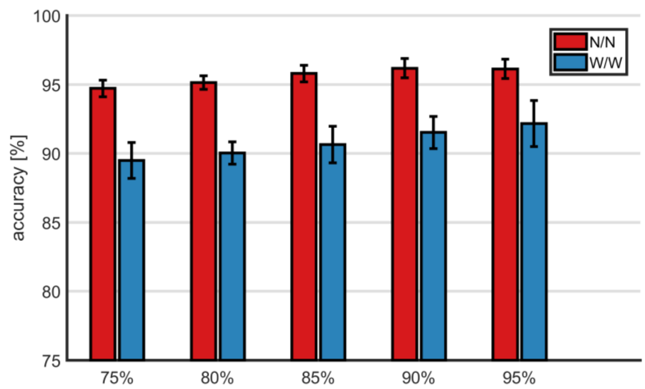


Fig. 6. Results of speaker identification (average accuracy with standard deviation) in dependence of percentage of database used for training for neutral (N/N) and whispered speech mode (W/W).

The results in N/W (neutral/whisper) train/test scenario for polynomial kernel and each speaker are presented in Table I. As can be seen, compared to N/N scenario the performance are degraded and very high difference between different speakers can be observed. Similar observation is found in whispering speaker identification with neutral trained HMM models [20].

TABLE I
ACCURACY FOR N/W SCENARIO (%)

Speaker	Accuracy
Speaker 1	65,2
Speaker 2	9,2
Speaker 3	21,4
Speaker 4	16,4
Speaker 5	62,0
Speaker 6	27,2
Speaker 7	25,4
Speaker 8	29,4
Speaker 9	33,4
Speaker 10	4,0
Average	**29,36**

V. CONCLUSION

Whispered speaker recognition is, by all means, a serious challenge for modern ASR systems.

As expected, whispering speaker ID was shown to be more difficult than recognition of a speaker that utters neutral speech. In this paper experiments on speaker identification in neutral and whisper mode for Whi-Spe speech database and SVM algorithm were conducted.

Future studies will be focused on examining more robust feature vectors in whispering speaker identification. Also, the application of the Teager energy operator [21-23] has shown improvements in robustness for whispered speech recognition, so there are reasonable expectations that it could help in speaker recognition as well.

As well, different machine learning algorithms are going to be analyzed (Gaussian mixture models and Neural networks).

ACKNOWLEDGEMENT

This research is supported by the project "Razvoj Internet of Things (IoT) aplikacija primjenom optičko-bežičnih tehnologija" of Ministry for Scientific and Technological Development, Higher Education and Information Society of Republic of Srpska.

REFERENCES

[1] J. Holms, W. Holms, *Speech Synthesis and Recognition*, Taylor & Francis, London, United Kingdom, 2001.

[2] C. Zhang, J. H. L. Hansen, “Analysis and classification of Speech Mode: Whisper through Shouted,” Proceedings of Interspeech 2007, pp. 2289-2292, 2007.

[3] T. Ito, K. Takeda, F. Itakura, “Analysis and recognition of whispered speech,” Speech Communication, vol. 45, no. 2, pp. 139–152, 2005.

[4] Y. Swerdlin, J. Smith, J. Wolfe, “The effect of whisper and creak vocal mechanisms on vocal tract resonances,” Journal of Acoustical Society of America 127, pp. 2590–2598, 2010.

[5] V. C. Tartter, “Identifiability of vowels and speakers from whispered syllables”, Perception & Psychophysics, vol. 49, no. 4, pp. 365–372, 1991.

[6] J. Galić, B. Popović, D. Šumarac Pavlović, “Whispered Speech Recognition using Hidden Markov Models and Support Vector Machines,” Acta Politechnica Hungarica, vol. 15, no. 5, pp. 11-29, 2018.

[7] J. Galić, S.T. Jovičić, Đ. Grozdić, B. Marković, “Constrained Lexicon Speaker Dependent Recognition of Whispered Speech,” Proceedings of International Symposium on Industrial Electronics and Applications (INDEL), pp. 180-184, 2014.

[8] A. Singh, A. M. Joshi, “Speaker Identification Through Natural and Whisper Speech Signal,” Optical and Wireless Technologies, Lecture Notes in Electrical Engineering, vol. 546, pp. 223-231, 2020.

[9] A. R. Naini, A. Rao, P. K. Ghosh, “Formant-gaps features for speaker verification using whispered speech,” IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, pp. 6231–6235, 2019.

[10] S. Furui, “50 Years of progress in speech and speaker recognition,” Proceeding of the International Conference Speech and Computer SPECOM, Patras, Greece, pp. 1-9, 2005.

[11] P. Clarkson, P. J. Moreno, “On the use of support vector machines for phonetic classification,” IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP 99, vol. 2, pp. 585-588, 1999.

[12] Z. Ćirković, Z. Banjac, “Jedna primena SVM klasifikatora u verifikaciji govornika nezavisno od teksta,“ Proceedings of the International Symposium Infoteh Jahorina, vol. 11, pp. 833-837, 2013.

[13] Z. Qu, Y. Liu, L. Zhang, M. Shao, “A Speech Recognition System Based on a Hybrid HMM/SVM Architecture,” First International Conference on Innovative Computing, Information and Control (ICICIC) Beijing, China, pp. 100-104, 2006.

[14] N. Kawaguchi, K. Takeda, S. Matsubara, I. Yokoo, T. Ito, K. Tatara, T. Shinde, F. Itakura, “Ciair speech corpus for real world applications,” International Joint Conference of the 5th Symposium on Natural Language Processing, pp. 288–295, 2002.

[15] P. X. Lee, D. Wee, H. S. Yin Toh, B. P. Lim, N. Chen, B. Ma, “Whispered Mandarin Corpus for Speech Technology Applications,” Proceedings of Interspeech 2014, pp. 1598–1602, 2014.

[16] T. Tran, S. Mariooryad, C. Busso, “Audiovisual corpus to analyze whisper speech,” presented at the International Conference on Acoustics, Speech and Signal Processing, pp. 8101–8105, 2013.

[17] F. Cummins, M. Grimaldi, T. Leonard, J. Simko, “The chains corpus: Characterizing individual speakers,” Proceedings of International Conference on Speech and Computer SPECOM, St. Petersburg, Russia, pp. 421–435, 2006.

[18] B. Marković, S. T. Jovičić, J. Galić, Đ. Grozdić, “Whispered speech database: design, processing and application,” In: Habernal, I., Matousek, V. (eds.), TSD 2013, LNAI 8082, Springer-Verlag Berlin Heidelberg, pp. 591–598, 2013.

[19] J. M. Garcia-Cabellos, C. Peleaz-Moreno, A. Gallardo-Antolin, F. Perez-Cruz, F. Diaz-de-Maria, “SVM classifiers for ASR: A discussion about paramterization,” Proceedings of 12th European Signal Processing Conference, pp. 2067-2070, 2004.

[20] X. Fan, J. H. L. Hansen, “Speaker identification within whispered speech audio stream,” IEEE Transactions on Audio, Speech and Language Processing, vol. 19, no. 5, 1408-1421, 2011.

[21] Đ. Grozdić,, “Application of Neural Networks in Whispered Speech Recognition”, PhD dissertation, University of Belgrade, Belgrade, Serbia, 2017.

[22] Đ. Grozdić, S. Jovičić, M. Subotić, “Whispered speech recognition using deep denoising autoencoder”, Engineering Applications of Artificial Intelligence, vol. 59, pp 15-22, 2017.

[23] Đ. Grozdić, S. Jovičić, “Whispered speech recognition using deep denoising autoencoder and inverse filtering”, IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 12, pp. 2313-2322, 2017.

Cepstrum-Based Pitch Detection of Industrial Product Sound

Dejan Ćirić, Marko Janković, Marko Milenković, Miljan Miletić

***Abstract*—Various audio features can be extracted from audio signals. One of very important ones is pitch. Different algorithms and methods have been proposed in literature to detect pitch. Among them, cepstrum-based pitch detection as a frequency domain method has often been used in practice. Cepstrum is calculated as the inverse Fourier transform of the logarithm of signal spectrum. The fundamental frequency and pitch in this way is estimated as the maximum value of cesptrum in the defined segment. Here, pitch of some industrial products (compressors and DC motors) are estimated by applying the modified cepstrum-based algorithm. The detected pitch values can be used to make a distinction between different working conditions of these products such as different rotation-per-minute (rpm).**

***Index Terms*—Cepstrum analysis; Pitch detection; Peak finding; Audio Feature; Audio signals; Industrial product sound.**

I. Introduction

PITCH detection is a common task present in a number of researches mostly related to speech, since the pitch (or fundamental frequency) is one of the most important parameters of speech. Thus, detection of pitch can be found in different speaker recognition and identification systems, speech synthesis systems, telecommunication systems, etc. [1-3]. In addition, pitch is one of the audio features (attributes) used for audio classification, detection and recognition by applying machine or deep learning [4].

Pitch can be detected in the time or frequency domain. One of simple time domain algorithms (methods) is the zero crossing rate method. The most important methods in the time domain are typically based on auto-correlation using a hypothesis that the auto-correlation function of a periodic signal is also periodic and that these two periods are coincident [5].

Regarding the frequency domain methods for pitch detection, one of the most popular is cepstrum-based method. Power cepstrum has some similar properties with the complex cepstrum obtained by homomorphic deconvolution [5], with the main difference that phase information is lost in the power cepstrum, which is called in the rest of the paper cepstrum.

Cepstrum, complex cepstrum, and homomorphic deconvolution have been applied in various areas such as audio processing, speech processing, geophysics, radar, medical imaging, etc. [5]. Some of the applications of both cepstrum and complex cepstrum include restoration of old phonograph recordings [6], cepstral pitch detector, speech recognition and speaker identification.

The periodicity present in a signal that is related to the pitch can be estimated from the cepstrum. Comparing with some other methods for pitch detection, the cepstrum method is able to provide accurate and robust results, but it is computationally complex [5].

This paper presents potentials for using pitch as an audio feature of sound of certain industrial products such as compressors and DC motors. The goal is to investigate if such a feature is able to provide clear distinction between different compressors or DC motors. The pitch is estimated by using cepstrum-based algorithm modified in a sense that peak finder is applied to the obtained cepstrum. Different compressors are related to compressors working with different rotations-per-minute (rpm), while different DC motors are related to different types of DC motors used in the automotive industry.

II. Pitch Detection Algorithms

Pitch is an important attribute of some audio signals such as speech signals. In speech, as a consequence of the vocal fold vibrations, the signal waveform contains certain periodicity translated into “pitch step” in the time domain and “fundamental frequency” or pitch in the frequency domain. However, pitch as an audio feature can also be of significance in machine and deep learning applied to a variety of audio signals including those containing sounds of industrial products, e.g. DC motors, home appliances or internal combustion engines of passenger vehicles [7].

There are various algorithms for pitch detection divided according to different criteria. Thus, there are block based and event based algorithms [9]. In the block based algorithms, the signal is sliced into small segments assuming that the pitch remains constant during the segments. On the other hand, event based algorithms use pitch marking or epoch detection. Here, pitch is not assumed to be constant over several pitch cycles. This is why these algorithms are able to track fast pitch changes even during the segments [10].

According to domain in which the algorithms are applied,

Dejan Ćirić is with the University of Niš, Faculty of Electronic Engineering in Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia (e-mail: dejan.ciric@elfak.ni.ac.rs).

Marko Janković is with the University of Niš, Faculty of Electronic Engineering in Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia (e-mail: marko.p.jankovic@elfak.rs).

Marko Milenković is with the Faculty of Arts, University of Niš, Kneginje Ljubice 10, 18105 Niš, Serbia (e-mail: lionheartmarko@gmail.com).

Miljan Miletić is with the Technical College of Applied Studies in Kragujevac, section in Kruševac, Kosančićeva 36, 37000 Kruševac, Serbia (e-mail: miljan.miletic@vhts.edu.rs).

they can be divided in three groups: time domain, frequency domain and hybrid group (combining time and frequency domain) [11]. Algorithms in the time domain use characteristics of a signal in the time domain, e.g. amplitude fluctuations, zero-crossing and auto-correlation attributes. This group contains the algorithms such as parallel processing time-domain method [12], data reduction method [13], modified auto-correlation (AUTOC), average magnitude difference function (AMDF) and YIN [14].

Frequency domain algorithms are based on the property that periodicity in the time domain is translated into series of peaks (impulses) in the frequency domain. This group of algorithms contains methods such as harmonic product spectrum (HPS), cepstrum-based pitch detection, linear predicting coding (LPC) and sawtooth waveform inspired pitch estimator (SWIPE) [15].

Hybrid detectors combine both the time and frequency domain algorithms. In that regard, frequency domain algorithms can yield temporary spectral aligned sound waves, and after that, auto-correlation methods are applied to determine the pitch period [11]. Hybrid pitch detection methods include pitch estimation filter with amplitude compression (PEFAC) [16], YAAPT [17], multi-band summary correlogram (MBSC) [18] and BaNa [19].

A. AUTOC as Time Domain Method

Among the time domain pitch detection methods, the most used one is the autocorrelation approach. It is based on finding the highest value of the auto-correlation function. Here, the auto-correlation function (*AF*) of a signal $s(n)$ (n=0, 1, … , N-1) is defined as

$$AF(k)=\frac{1}{N}\sum_{n=0}^{N-1-k} s(n)s(n+k),\quad k=0,1,...,N-1, \tag{1}$$

where N is the signal (or frame) length, while k is the lag index. The pitch is detected at the location of the peak of auto-correlation function.

B. Cepstrum-based Algorithm as Frequency Domain Method

Cepstrum $C(m)$ can be calculated as the inverse Fourier transform of the logarithm of Fourier transform of the target signal, $s(n)$:

$$\begin{gathered} S(k)=\log\left\{\sum_{n=0}^{N-1} s(n)\cdot\exp^{-j\frac{2\pi}{N}nk}\right\} \\ C(m)=\frac{1}{N}\left\{\sum_{k=0}^{N-1} S(k)\cdot\exp^{j\frac{2\pi}{N}mk}\right\} \end{gathered}. \tag{2}$$

The pitch is detected at the location of cepstrum maximum calculated as given in (2). It is presented in the literature that the cepstrum-based method is sensitive to noise in the target signal [11].

C. PEFAC as Hybrid Method

The pitch in the PEFAC algorithm is detected by convolving the power spectral density of the signal in the log-frequency domain with the filter summing the energy of the pitch harmonics [11]. The model at the time moment t of a perfectly periodic signal (having fundamental frequency f_0) in the power spectral density domain can be expressed as

$$Y_t(f)=\sum_{k=1}^{K} a_{k,t}\delta(f-kf_0)+N_t(f), \tag{3}$$

where $N_k(f)$ is the power spectral density of the undesired noise and $a_{k,t}$ is the power of the k-th harmonic. The signal model gets the following form in the logarithmic domain

$$Y_t(f)=\sum_{k=1}^{K} a_{k,t}\delta(q-\log(k)-\log(f_0))+N_t(f), \tag{4}$$

where $q = \log(f)$. The energies of the signal components in this domain can be combined by convolving $Y_t(q)$ with the impulse response filter.

The filter $h(q)$ can suppress the noise with smoothly varying spectra, but this is not the case for high amplitude narrowband noise. This is why the spectrum compression is applied before the convolution with the filter $h(q)$

$$Y_t'(q)=Y_t(q)^{a_t(q)} \tag{5}$$

where t represents time index and $a_t(q)$ represents the compression exponent [11].

III. Methods Applied

For the purpose of carrying out this research, audio signals with different spectral contents are selected. Some of them have known pitch and harmonics distribution, such as the trumpet sound given in Fig. 1, while in the other signals (containing the sounds of certain industrial products) these parameters are unknown. The sound of trumpet is tonal sound with pronounced pitch, and the main characteristic of such a sound is periodicity, which can be seen in the time domain as presented in Fig. 2.

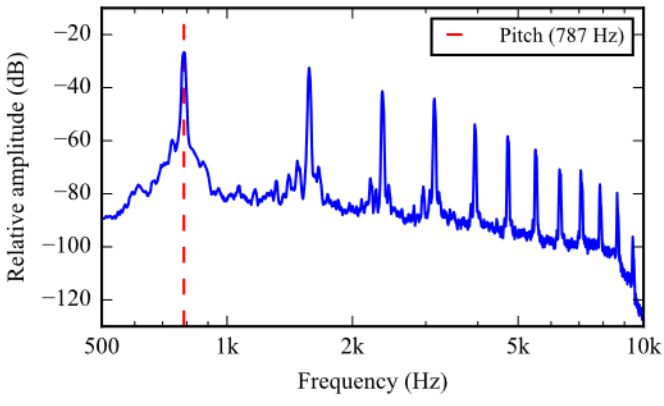


Fig. 1. Spectrum of trumpet sound consisting of fundamental (f_0) at frequency of 787 Hz and partials that have a harmonic distribution relative to the fundamental.

Cepstrum-based pitch detection uses signal periodicity. In this case, the periodicity refers to a particular waveform of the specific length that is repeated throughout the signal, and it is reflected in a discrete spectrum containing prominent peaks equally distributed throughout the frequency range.

The cepstrum-based pitch detection algorithm consists of five main steps described below. In the first step, the signal is transformed from the time to the frequency (spectral) domain using the Fourier transform. By applying the relevant function in Python, two variables are obtained - spectrum (given in complex numbers) and frequency vector. In the second step, the logarithm of the spectrum magnitude is calculated. The basic idea of cepstrum is to transfer the periodicity from the logarithmic representation of the spectrum to the time domain. For that reason, in the third step, the inverse Fourier transform is applied over the logarithm of the spectrum magnitude data. The cepstrum of the trumpet sound from Figs. 1 and 2 obtained in the described way is shown in Fig. 3.

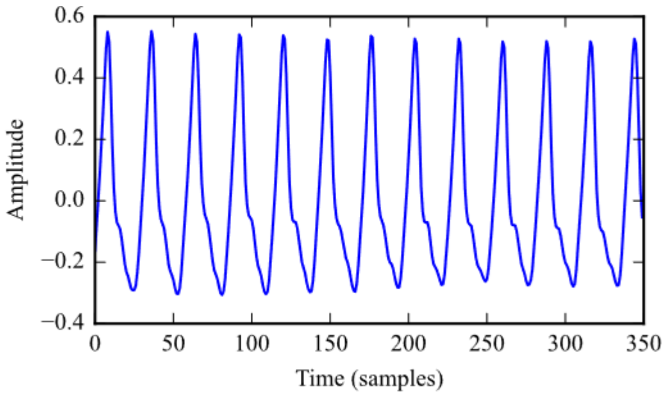


Fig. 2. Trumpet sound shown in time domain.

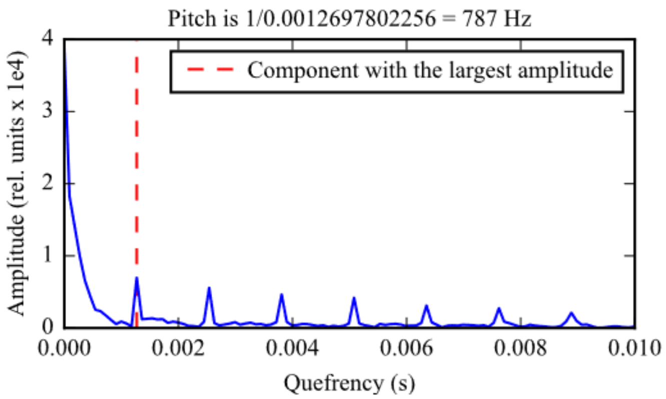


Fig. 3. Cepsrum of trumpet sound.

As can be seen in Fig. 3, the *x*-axis is titled quefrency, which is an anagram of frequency and it is related to time scale. The reason is in performing the inverse Fourier transform, that is, in inverting the frequency by this transformation. As a consequence, the highest frequencies are located at the beginning of the *x*-axis, while the lowest frequencies are located at the end of the *x*-axis. Frequencies are converted into quefrencies by taking 1/frequency.

The fourth step is related to finding the maximum value of the cesptrum, that is, finding the quefrency value at which the cepstrum maximum occurs. In majority of cases, the maximum amplitude of the cepstrum is located at the beginning of the x-axis – at the zero quefrency or its vicinity, which greatly complicates the pitch detection in an automated manner. There are several approaches to overcome this problem. One of them is to limit the quefrency band in which the search of cepstrum maximum is carried out. In this way, an error can be introduced in estimating the pitch of some sounds, such as sound of certain industrial products, since it is not known in advance in which frequency band to expect the pitch. The second approach is not to detect the pitch at the absolute maximum of the cepstrum, but instead to skip the cepstrum maximum at the zero quefrency, and to use the second largest value of the cepstrum for the pitch estimation. Such an approach leads to an error of pitch detection in the case where the cepstrum do not have the maximum value at the zero quefrency.

In this paper, an alternative approach to solve the mentioned problem is applied. Thus, the function for finding peaks (*find_peaks*) from the library *scipy* in *Python* is used. This function finds all local maxima by simple comparison of neighboring values, with the ability to define large number conditions for the peak properties. Due to the fact that there are no neighboring values on the left side of the cepstrum maximum at the zero quefrency (or in close vicinity), this maximum is automatically not considered as a peak.

When the relevant maximum of the cepstrum is selected, the quefrency value of that maximum is converted into frequency representing the estimated pitch (pitch = 1/(quefrency of cepstrum maximum)).

To check the described algorithm, the fundamental frequency from the trumpet sound is removed by filtering, see the spectrum shown in Fig. 4.a). The determined maximum of the cepstrum is located at the same quefrency position, see Fig. 4.b), as before removing the fundamental component.

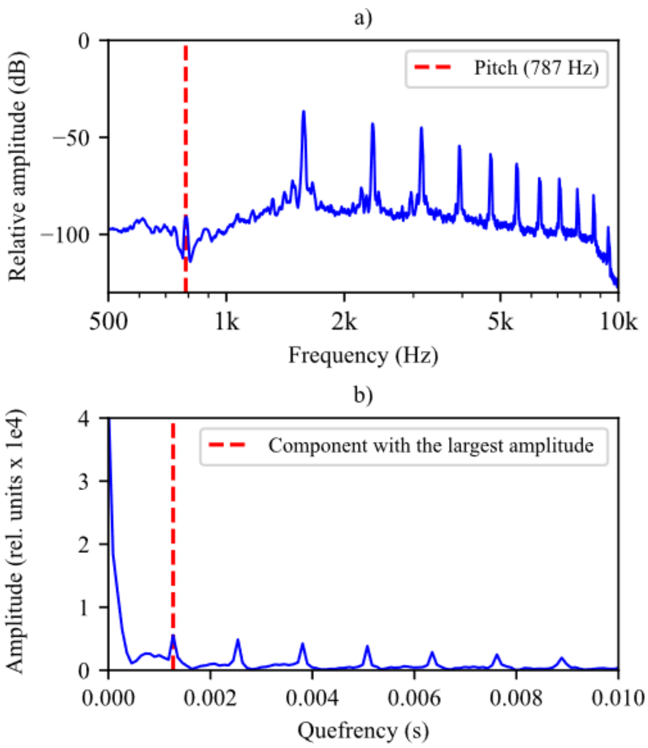


Fig. 4. Spectrum a) and cepstrum b) of trumpet sound after removing the fundamental component located at 787 Hz.

For the purpose of this research, the estimation of pitch by the modified cepstrum-based algorithm described above is done on different sounds of industrial products including fridge compressors and DC motors. The sound of fridge compressors were recorded in the semi-anechoic chamber in three working cycles (modes of operation) having different rpms (4000 rpm, 2400 rpm and 1300 rpm). Sounds of DC motors were also recorded in the semi-anechoic chamber (not the same one used for the compressor recording) within the production line. The recording was done on two different types of DC motors (here denoted type A and type B), in two different directions of rotation (here denoted direction 1 and direction 2) and for two different conditions regarding the failure (without failure and with certain failure). The analysis of recorded audio signals is carried out using the scripts developed in Python 3.8.

IV. Analysis of Detected Pitch

Th
e potentials for applying the cepstrum-based pitch detection in making difference between different samples or working conditions of certain industrial products is investigated here focusing on spectrum and cepstrum of the sounds of these products, that is, on one-figure value of the detected pitch.

The first product whose sound is analyzed is fridge compressors. The target for this product is to consider if it is possible to make a distinction between three different working conditions of compressors – three different rpms (4000, 2400 and 1300). The spectrum and cepstrum of the compressors having 4000 rpm are shown in Fig. 5. The pitch estimated by the described procedure is 6400.05 Hz.

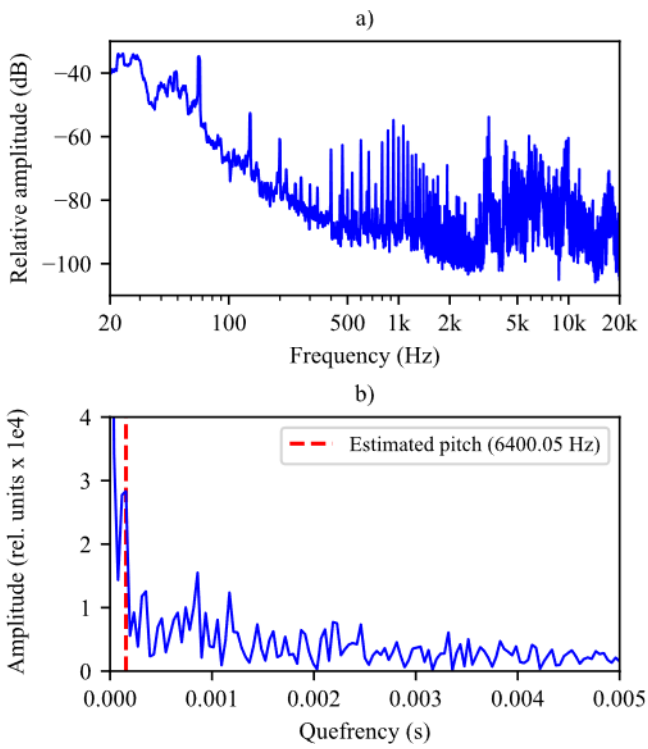


Fig. 5. Spectrum a) and cepstrum b) of fridge compressor sound at 4000 rpm.

When the rpm is changed from 4000 to 2400, it causes certain changes in the spectrum, but also in the cepstrum and consequently in the detected pitch. Fig. 6 shows the spectrum and cepstrum of the fridge compressor sound at 2400 rpm, where the estimated pitch is 2756.25 Hz.

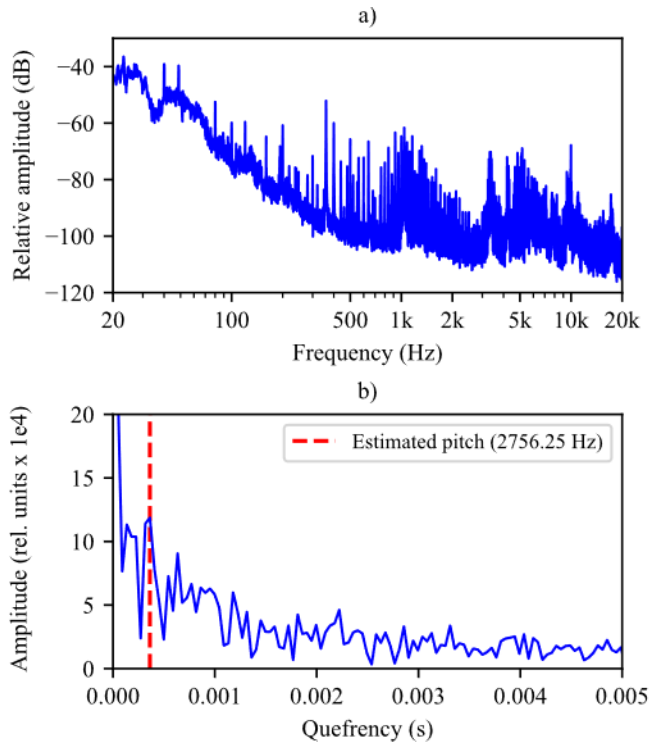


Fig. 6. Spectrum a) and cepstrum b) of fridge compressor sound at 2400 rpm.

By changing the rpm to 1300, the detected pitch is changed to 5120.04 Hz. The spectrum and cepstrum for that rpm are presented in Fig. 7. Regarding the cepstrum, not only the relevant maximum of the cepstrum, but also its pattern is changed by changing rpm. This is why it seems reasonable to introduce at least one more attribute that will reflect dissimilarity of the cepstrum pattern. This attribute can be related either to decay of the cepstrum envelope, cepstrum energy within certain qefrency limits or even a vector containing the cepstrum values within pre-defined qefrency limits.

The next step in the analysis includes DC motors. Spectrum and cepstrum of the same DC motor of type A, but in two opposite directions of rotation are shown in Fig. 8. Comparing the spectra, it can be concluded that overall pattern is similar for both directions of rotation, with certain differences in particular frequency bands and at particular frequencies. The patterns of cepstrum are also similar, but still having some differences for different directions of rotation. The estimated pitch for both directions is the same, 2666.33 Hz, since the maximum value of cepstrum is located at the same quefrency.

Comparison of spectra and “cepstra” of sounds of two DC motors of different types (A and B) is presented in Fig. 9. Shapes of the spectra given in Fig. 9.a) and Fig. 8.a) for the motor A are similar, with certain differences, since different motors of the same type A are used for the analysis. The same situation exists in the “cepstra” from Fig. 8.b) and Fig. 9.b). In spite of these differences, the estimated pitch for two motors of the same type A is the same (2666.73 Hz), while the estimated pitch for the motor of type B is 888.91 Hz.

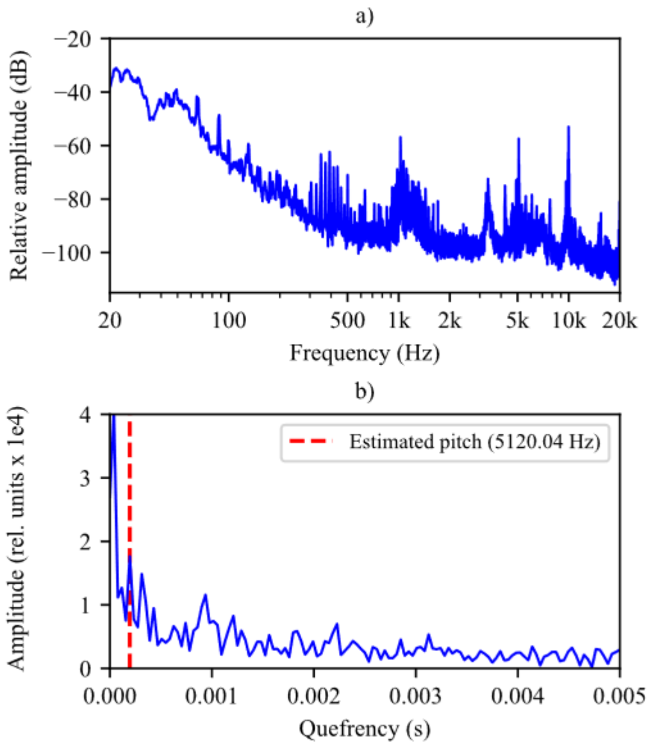


Fig. 7. Spectrum a) and cepstrum b) of fridge compressor sound at 1300 rpm.

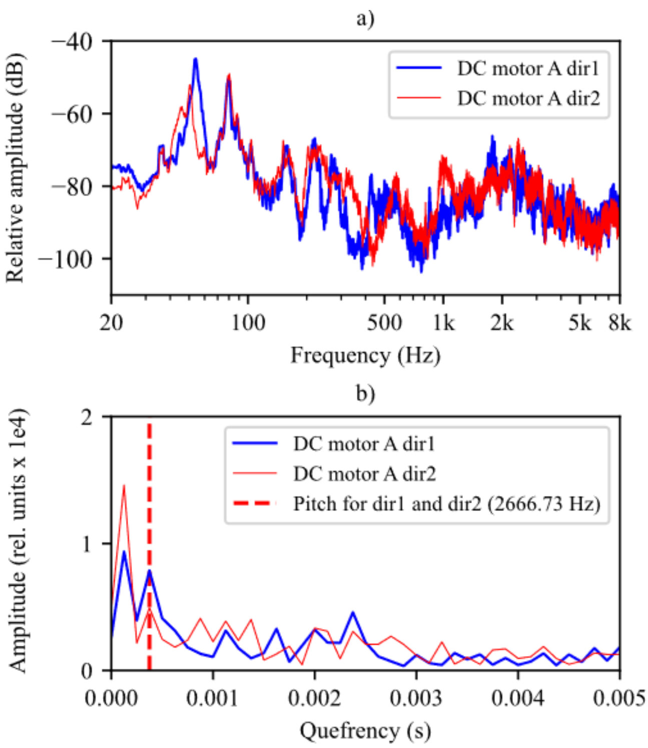


Fig. 8. Spectrum a) and cepstrum b) of DC motor A in two opposite directions of rotation.

When detected pitch values for DC motors with and without failures are analyzed, the differences between motors depend on the failure itself. In some cases, the failure causes change of periodicity or pseudo-periodicity of sound waveform leading to a certain change of the estimated pitch. Fig. 10 illustrates one of such cases presenting spectra and “cepstra” of DC motors without failure (OK motor) and with failure (NOT OK motor). The estimated pitch for OK motor is 2666.73 Hz and for NOT OK motor is 1333.36 Hz. Here, the differences between the spectra and “cepstra” are larger in comparison to the previously analyzed two cases of DC motors.

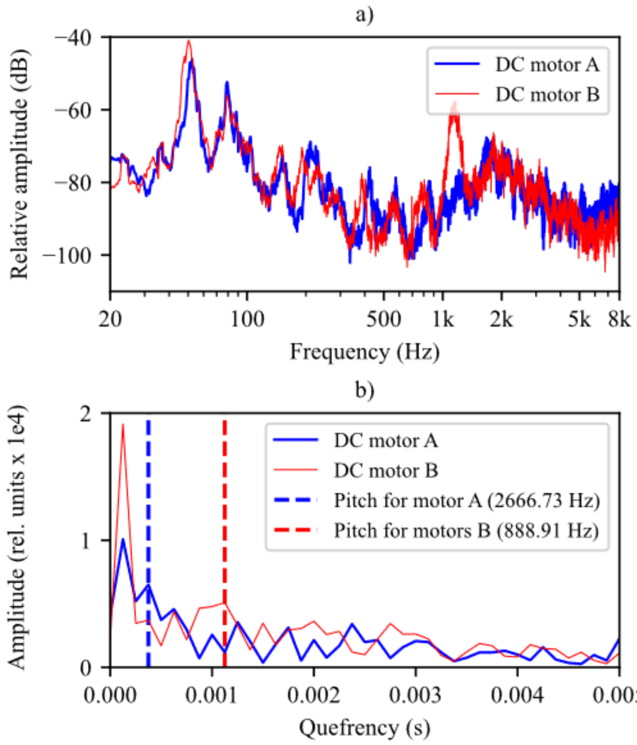


Fig. 9. Spectrum a) and cepstrum b) of different DC motors (motors of type A and B) and direction of rotation 1.

V. Conclusion

This paper analyses potentials for making a difference between working conditions or states of two industrial products, fridge compressors and DC motors, based on pitch of their sounds. The pitch is detected using the cepstrum-based algorithm.

The results show that there are conditions and states where the estimated pitches are different for different conditions (states). However, there are also cases where the pitch only could not be used for making a difference between conditions (states) of these products. Even in such cases, the patterns or shapes of the “cepstra” show certain differences for different conditions (states). This is why a measure calculated from the cepstrum can be introduced in addition to pitch that can be used as a new audio feature.

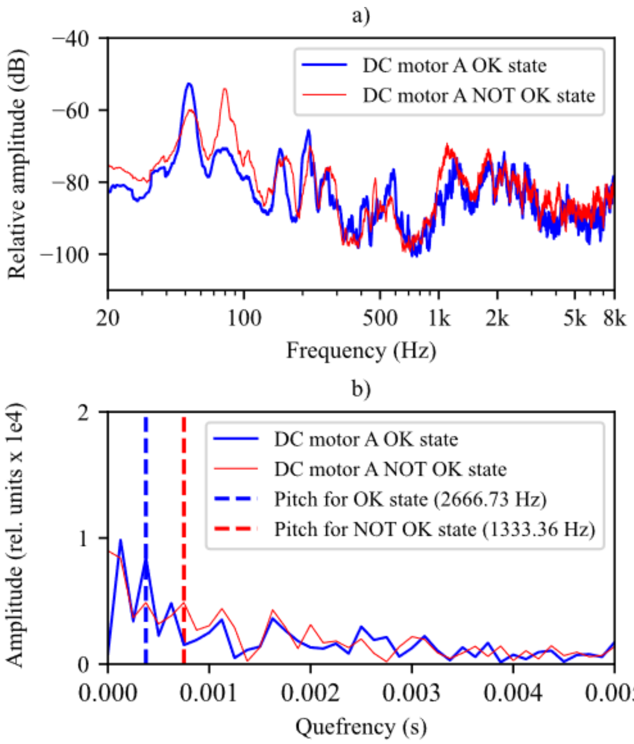


Fig. 10. Spectrum a) and cepstrum b) of OK DC motor (without failure) and NOT OK DC motor (with failure).

ACKNOWLEDGMENT

This research was supported by the Science Fund of the Republic of Serbia, 6527104, AI-Com-in-AI.

REFERENCES

[1] S. Makov, A. Minaev, I. Grinev, D. Cherhyshov, B. Kudruavcev, V.Mladenovic, "A spectral-based pitch detection method," Proc. AIP Conference (XV International Scientific-Technical Conference "Dynamic of Technical Systems"), Rostov-on-Don, Russia, 2188, 050005, 11-13 September, 2019.

[2] B. Gold, N. Morgan, D. Ellis, *Speech and Audio Signal Processing: Processing and Perception of Speech and Music*, 2nd ed. USA: Willey, 2011.

[3] X. Zhao, Y. Shao, D. Wang, "CASA-Based robust speaker identification," *IEEE Transactions on Audio Speech and Language Processing*, vol. *20*, no. *5*, pp. *1608-1616*, July, 2012.

[4] A. B. Nielsen, L. K. Hansen, U. Kjems, "Pitch based sound classification," Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toulouse, France, vol. 3, 14-19 May, 2006.

[5] K. Pratibha, H.M. Chandrashekar, "Estimation and tracking of pitch for noisy speech signals using EMD based autocorrelation function algorithm," Proc. 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, pp. 2071-2075, 19/20 May, 2017.

[6] T.G. Stockham, Jr., T.M. Cannon, R.B. Ingebretsen, "Blind deconvolution through digital signal processing," *Proc. IEEE*, vol. 63, pp. *678-692*, April, 1975.

[7] M. Milivojčević, D. Ćirić, "Extracting relevant acoustic features of the internal combustion engines," Proc. 7th International Conference on Electrical, Electronic and Computing Engineering (IcETRAN), Belgrade, Čačak, Niš, Novi Sad, Serbia, pp. 49-53, 28-29 September, 2020.

[8] A. V. Oppenheim, R. W. Schafer, "From frequency to quefrency: A history of the cepstrum," *IEEE Signal Processing Magazine*, pp. *95-106*, September, 2004.

[9] H. Chander, B. Singh, R. Khanna, "Short time cepstrum analysis method for pitch estimation of an arbitrary speech signal using MATLAB," *International Journal of Research in Electronics and Computer Engineering*, vol. *5*, no. *3*, pp. *1-4*, July-September, 2017.

[10] T. Ananthapadmanabha, B. Yegnanarayana, "Epoch extraction of voiced speech," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. *ASSP-23*, no. *6*, pp. *562-575*, December, 1975.

[11] L Sukhostat, Y. Imamverdiyev, "A comparative analysis of pitch detection methods under the influence of different noise conditions," *Journal of Voice*, vol. *29*, no. *4*, pp. *410-417*, July, 2015.

[12] B. Gold, L. Rabiner, "Parallel processing techniques for estimating pitch periods of speech in the time domain," *Journal of the Acoustical Society of America*, vol. *46*, no. *2*, pp. *442-448*, August, 1969.

[13] N. J. Miller, "Pitch detection by data reduction," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. *23*, no. *1*, pp. *72-79*, February, 1975.

[14] A. de Cheveigne, H. Kawahara, "YIN, a fundamental frequency estimator for speech and music," *Journal of the Acoustical Society of America*, vol. *111*, no. *4*, pp. *1917-1930*, April, 2002.

[15] A. Camacho, "SWIPE: A sawtooth waveform inspired pilch estimator for speech and music," Ph.D. dissertation, University of Florida, Gainesville, Florida, USA, 2007.

[16] S. Gonzalez, M. Brookes, "A pitch estimation filter robust to high levels of noise (PEFAC)," Proc. 19'th European Signal Processing Conference - EUSIPCO, Barcelona, Spain, pp. 451-455, 29 August – 02 September, 2011.

[17] K. Kasi, S. A. Zahorian, "Yet another algorithm for pitch tracking," Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing - ICASSP, Orlando, Florida, USA, 13-17 May, 2002.

[18] L. N. Tan, A. Alwan, "Multi-band summary correlogram-based pitch detection for noisy speech," *Speech Communication*, vol. *55*, pp. *841-856*, 2013.

[19] H. Ba, N. Yang, I. Demirkol, W. Heinzelman, "BaNa: a hybrid approach for noise resilient pitch detection," Proc. IEEE Statistical Signal Processing Workshop (SSP), Ann Arbor, Michigan, USA, 04-07 August, 2012..

АНТЕНЕ И ПРОСТИРАЊЕ
/
ANTENNAS AND PROPAGATION

(АП/API)

Karakteristike materijala za štampane antene u opsegu 65-110 GHz

Nikola Bošković i Miloš Radovanović

Apstrakt—**Izbor materijala igra presudnu ulogu u karakteristikama štampanih antena. Ovo je naročito kritično na W- opsegu, gde je mehanizam gubitaka značajno izmenjen, zbog čega veoma mali broj standardnih dielektrika može biti korišćen. Radi suzbijanja parazitnih modova i površinskih talasa, debljina dielektrika mora biti veoma mala u odnosu na učestanost, što dovodi do toga da efekat površinske hrapavosti bakarne provodne folije ima dominantan uticaj na gubitke. U ovom radu je prikazan značaj izbora odgovarajuće folije kao i njen uticaj. Zaključci su praćeni simulacijama i merenjima.**

Ključne reči—**Balansni mikrostrip, gubici u materijalu, štampana tehnologija, štampane antene.**

I. Uvod

Štampane antene su klasa antena koje se tipično sastoji od provodnika koji je najčešće u vidu tanke bakarne folije, i dielektrika na kome se provodna folija nalazi. Sama antena se može sastojati od više slojeva, kako dielektrika tako i provodnika. Neke od primarnih karakteristika štampanih antena su mala cena, mogućnost masovne proizvodnje, lako formiranje antenskih nizova u više prostornih dimenzija. Sa antenskim nizovima moguće je postići veliki dobitak, i proizvoljnu širinu glavnog snopa zračenja.

Izbor materijala na kome će biti napravljena antena ima presudan uticaj na fundamentalne parametre antene: radni opseg, učestanost, dobitak, dimenzije. Veliki uticaj materijala je svakako na efikasnost antene, tj. gubitke. Gubici zavise od mnogih faktora, i menjaju se sa frekvencijom tako da je željena radna učestanost početni uslov za izbor odgovarajućeg materijala.

Gubici u mikrostripu mogu tipično biti podeljeni na: gubitke u dielektriku, gubitke u metalu, gubitke usled neželjenog zračenja i gubitke usled procesa izrade (tolerancija). Gubici usled neželjenog zračenja mogu biti izuzetno kompleksni, jer mogu poticati od više različitih izvora: parazitni modovi, površinski talasi, zračenje od strane napojne mreže, itd. Za antenu ovo može biti naročito problematično, jer pored smanjenja efikasnosti može prilično degradirati dijagram zračenja antene. Da bi se minimizovalo neželjeno zračenje opšti princip je korišćenje tankog dielektrika u odnosu na radnu učestanost i uskih vodova. U tankim dielektricima sa malim gubicima izraženim preko tanδ koji je definisan kao odnos imaginarnog i realnog dela kompleksne permitivnosti (1), gubici u metalu postaju dominantni.

$$\varepsilon_r = \varepsilon_r' - j\varepsilon_r'' = \varepsilon_r'(1-\tan\delta) \qquad (1)$$

Na učestanostima oko 2.4 GHz i niže pretežno se koristi FR-4, zbog veoma niske cene i lake dostupnosti, na višim učestanostima koristi se znatno skuplji dielektrici poput Rogers 3000, 4000 ili 5000 serije. Na nižim učestanostima gubici u dielektriku su dominantni, pa se zbog toga gubici u dielektriku uzimaju kao primaran parametar za računanje gubitaka. Kod štampanih antena dielektrik pored toga što razdvaja provodne delove antene, se ponaša kao nosač provodnika. Na GSM opsegu se mogu koristi znatno deblje folije od tipičnih 17 mikrona. U ovom slučaju kao dielektrik se može koristiti vazduh, jer se debela folija neće deformisati pod svojom težinom. Vazduh ima dielektrične osobine slične vakuumu. Sa tanjom folijom kao dielektrične podloge moguće je koristi penaste materijale poput Rohacell-a koji imaju malu dielektričnu konstantu $\varepsilon_r \approx 1.05$ (tanδ ≈ 0.0002).

Debljina folije koja će biti korišćena je određena na osnovu skin efekta, po kome će se struja sa rastom učestanosti koncentrisati u veoma tankom sloju po površini provodnika, [1]. Ovo je osnovni razlog zašto je moguće koristi veoma tanke folije kao provodnike na visokim učestanostima. Drugi veoma bitan razlog za korišćenje tankih folija na visokim učestanostima je površinska hrapavost folije, koja je ovde primaran uzrok gubitaka. Gubici u metalu tj. foliji postaju veoma izraženi i za dielektrik sa niskim tanδ postaju dominantni. Deblja folija će tipično imati veću hrapavost.

Dielektrična konstanta kao elementarna osobina dielektrika, primarno utiče na dimenzije antene. Za veći ε_r dimenzije antene će biti manje, ali takođe efekat površinskih talasa i neželjenog zračenja biće znatno više izražen. Vakuum ili vazduh bi za mnoge slučajeve bio idealan dielektrik. ε_r je disperzivan parametar. Za komercijalne dielektrike ε_r je tipično data kao prosečna numerička vrednost na 10 GHz, dobijena sa serije uzoraka korišćenjem određene metode, [2].

Kod mikrostripa linije električnog polja antene putuju između dva medijuma: dielektrika i vazduha. Za analizu se koristi koncept efektivnog dielektrika dielektrične konstante ε_{eff}. Po tome umesto nehomogene sredine uzima se da linije električnog polja prostiru kroz ekvivalentni homogeni dielektrik. ε_{eff} ima ogromni značaj, jer se može direktno odrediti merenjima, a relacija (2) između ε_{eff} i ε_r je dobro poznata, [1],

Nikola Bošković – Institut za fiziku, Univerzitet u Beogradu, Pregrevica 118, 11080 Zemun, Srbija (e-mail: nikolab@ipb.ac.rs).
Miloš Radovanović – Institut za fiziku, Univerzitet u Beogradu, Pregrevica 118, 11080 Zemun, Srbija (e-mail: rmilos@ipb.ac.rs).

$$\varepsilon_{\text{eff}} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2} \frac{1}{\sqrt{1 + 12\frac{h}{w}}} \tag{2}$$

gde su h i w, debljina dielektrika i širina mikrostrip voda. Postoje još preciznije relacije koje uzimaju u obzir debljinu bakarne folije. Svaka relacije ima određenu grešku kao i uslove pod kojima je primenjiva. U ovom radu precizna korelacija između ε_{eff} i ε_r je određena na osnovu simulacionog modela.

Površinska hrapavost folije ne utiče samo na povećanje gubitaka u metalu sa rastom učestanosti, već takođe utiče na promenu faze prostiranja talasa, tj. utiče na promenu ε_{eff}, tako da na osnovu toga dva identična dielektrika sa različitim profilima bakarne folije mogu imati dosta drugačiju ε_{eff}, [3].

U ovom radu je prikazana širokopojasna analiza dielektrika Rogers 3003, debljine $h = 0.127$ mm, $\tan\delta = 0.001$, $\varepsilon_r \approx 3$.

II. ANALIZA I MERENJA

Jedan od najbitnijih faktora za dobijanje prototipa štampane antene koji se slaže sa proračunima i simulacijama je korišćenje tačnih parametara materijala. Ovo u praksi znači da korišćenjem parametara definisanih od strane proizvođača koji u većoj ili manjoj meri mogu odstupati od konkretnog primerka dielektrika, u startu unosi određenu grešku. Takođe različite metode za merenje dielektrične konstante neće nužno dati identične rezultate, [2]. Dobra praksa je da se pre izrade štampane antene izvrši procena dielektrične konstante na opsegu od interesa korišćenjem identičnog proizvodnog procesa kao za izradu same antene.

Za procenu dielektrične konstante u ovom radu korišćena je metoda faznih razlika, [4]. Ova metoda podrazumeva korišćenje vodova različitih dužina i poznatom razlikom u dužinama. Na osnovu merenja S-parametara oba voda, dobija se korelacija (3), između ε_{eff}, električne (ΔEL) i fizičke dužine (Δl) na posmatranoj frekvenciji f kao

$$\varepsilon_{\text{eff}} = \left(\frac{\Delta EL \cdot c}{360\,° \cdot f \cdot \Delta l} \right)^2 \tag{3}$$

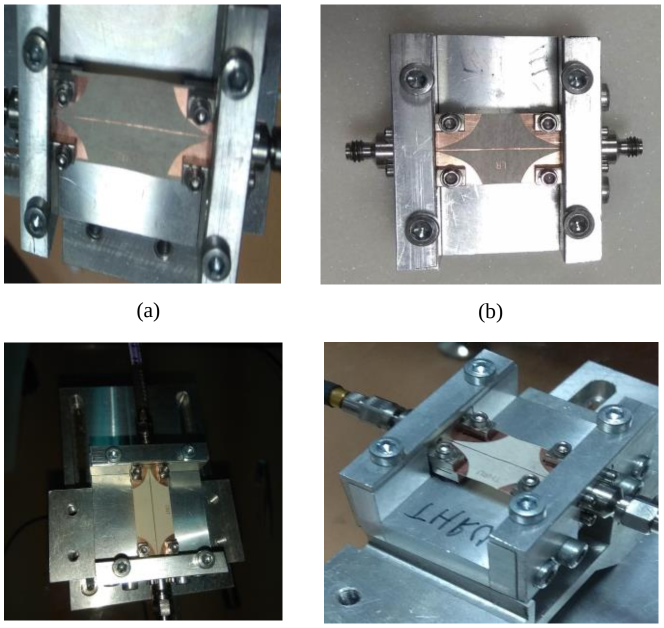
gde je c brzina svetlosti u vakuumu. Fizičke dimenzije korišćenih vodova su: širina voda $w = 0.2$ mm, dužine vodova su 20 mm i 30 mm, tj. razlika je $\Delta l = 10$ mm. Za merenje je korišćen Anritsu VectorStar ME7838A VNA. Meren je opseg od 65-110 GHz, sa korakom od 10 MHz. Konektori su kalibrisani sa Triple-Offset-Short-Through (SSST) kalibracijom. Preciznost izrade vodova je oko 0.01 mm. Dimenzije vodova su merene pod mikroskopom.

Velika prednost ove metode je što se vrlo lako može dobiti širokopojasna karakteristika ε_r. Tačnost ove metode veoma zavisi od tačnosti izmerene razlike faza od datih vodova. Potrebno je koristi veoma kvalitetne širokopojasne konektore, bez lemljenja. Ovde su korišćeni Southwest Microwave 1 mm koaksijalni konektori za W-opseg, deklarisani za rad od 0 do 110 GHz.

Sl. 1. Balansni mikrostrip vod od bakarne folije na dilektriku.

U datom slučaju vršena su merenja i procena za dva različita slučaja bakarne folije. Za slučaj ED (elektro-deponovanog) bakra i za slučaj rolovane bakarne folije. Od interesa je utvrđivanje parametara za balansni mikrostrip [5], Sl. 1. Kao što se vidi za razliku od mikrostripa kod koga je jedan vod (masa) znatno širi od drugog voda, kod balansnog mikrostripa oba voda su iste širine. Ovo u praksi znači da će balansni vod biti delikatniji za fizičku obradu, jer oba voda moraju da se fizički obrade na odgovarajuće dimenzije.

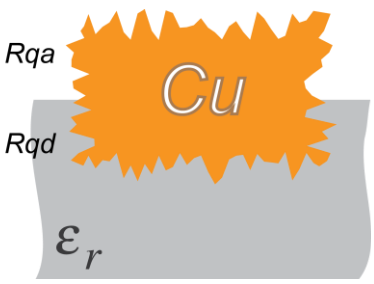
Aparatura za merenje je prikazana na Sl. 2. Kako je dielektrik veoma tanak, napravljeni su metalni nosači koji drže dielektrik i konektore tokom merenja.



(a) (b)

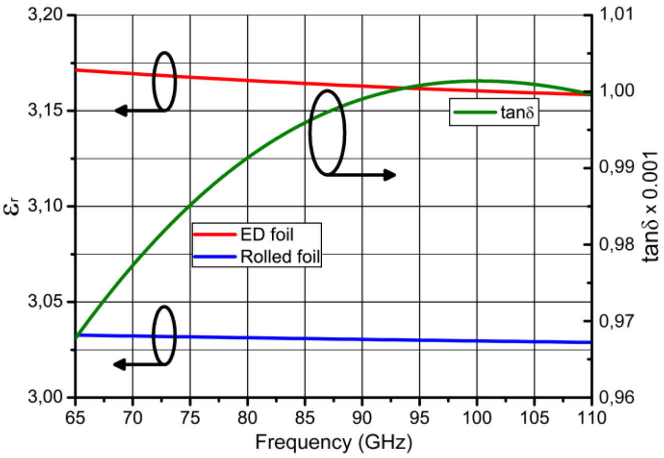
(c) (d)

Sl. 2. Vodovi različitih dužina na aparaturi za merenje S-parametara. Modeli sa rolovanom folijom (a-b) i ED folijom (c-d). Duži vodovi su dati pod (a) i (c).



Sl. 3. Prikaz poprečnog preseka bakarnog voda uronjenog u dielektrik sa prikazanom hrapavosti između vazduha i folije (*Rqa*) i hrapavosti između folije i dielektrika (*Rqd*).

Hrapavost se tipično predstavlja preko srednjeg kvadratnog korena hrapavosti, *Rq* [6]. Za slučaj bakarne folije za dielektrike daju se dve, tipično različite vrednosti. Jedna za površinu okrenutu ka vazduhu (*Rqa*), a druga za površinu okrenutu ka dielektriku (*Rqd*), Sl. 3. *Rqd* će uglavnom imati značajno veću vrednost i za neke slučajeve se čak veštački uvećava kako bi se omogućila što čvršća veza između folije i dielektrika. U ovom slučaju za rolovanu foliju vrednosti *Rqd* i *Rqa* su 0.4 μm i 0.3 μm, dok za ED vrednosti su redom 2 μm i 0.4 μm. Simulirana širokopojasna karakteristika tanδ je dobijena na osnovu Debajevog modela višeg reda [7]. U korišćenom modelu uticaj tanδ je značajano manji u odnosu na uticaj gubitaka u metalu.



Sl. 4. Vrednosti ε_r dobijene na osnovu merenja i aproksimacija tanδ.

Sa Sl. 4 se vidi da na osnovu merenja ε_{eff} odgovarajuće vrednosti ε_r se značajno razlikuju, iako je u pitanju isti dielektrik. Takođe se vidi da za slučaj rolovane bakarne folije dielektrična konstanta je veoma slična deklarisanim vrednostima na 10 GHz, gde je uticaj hrapavosti daleko manji. Sa grubljim profilima bakarne folije razlika postaje značajna [3]. Kako su grublji profili daleko jeftiniji njihova upotreba je znatno veća. Samim tim je jasno da u većini slučajeva deklarisani parametri dielektrika neće dati realističan elektromagnetni model, već vrednosti dobijene na ovaj način treba da budu ulazni parametri odgovarajućeg modela.

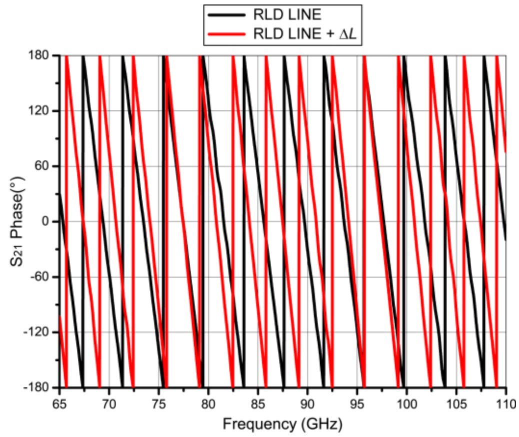
Druga velika razlika će biti naravno u pogledu gubitaka. Totalni gubici po jedinici dužine se mogu dobiti takođe merenjem prototipa sa Sl. 2, iz vrednosti magnitude S_{12}. Za uračunanje gubitka u metalu često se koristi koncept efektivne provodnosti metala σ_{eff}, po kome se umesto 58 MS/m za provodnost bakra, uzima manja vrednost dok gubici u modelu ne postanu jednaki stvarnim gubicima. Ovaj model je odličan za brze proračune, međutim on je veoma uskopojasan. Postoji veliki broj znatno naprednijih modela, od kojih većina zahteva veliki broj parametara za kvalitetnu procenu.

Uporedo sa empirijskim razvijani su i simulacioni 3D modeli. Ono što je posebno problematično kod ovog pristupa je što zbog veoma malih dimenzija hrapavosti, čak i veoma mali modeli dimenzija 1 mm bi imali ogroman broj nepoznatih. Takođe i postavka samih modela se često zasniva na posmatranju površine kao fraktalne strukture. U praksi hrapavost je neuređena tj. svaki pik ili dolina imaće nešto drugačiju vrednost. Za primenu u polju mikrotalasne tehnike potreban je model, koji može na osnovu dostupnih parametra, da precizno proceni uticaj hrapavosti na prostiranje elektromagnetskih talasa.

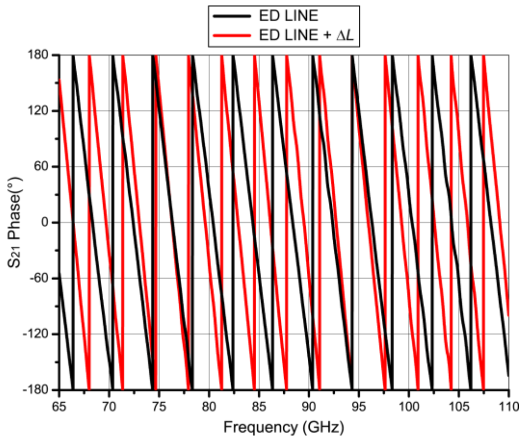
Posebno zgodan je gradijentni model [8], po kome je moguće dobiti širokopojasnu procenu uticaja hrapavosti samo na osnovu *Rq*. Osnova ovog modela je da se umesto konstante vrednosti σ_{eff}, uzima vrednost dobijena na osnovu statističke kumulativne funkcije raspodele kao (4)

$$\sigma(x) = \sigma_{Cu} \int_{-\infty}^{x} \exp\left(-\frac{u^2}{2R_q^2}\right) du \qquad (4)$$

σ se posmatra kao funkcija rastojanja x u pravcu normalnom na površinu gde se nalazi bakarna folija. U ovom opsegu funkcija može imati sve moguće vrednosti od nule do provodnosti bakra bez hrapavosti σ_{Cu}.



(a)

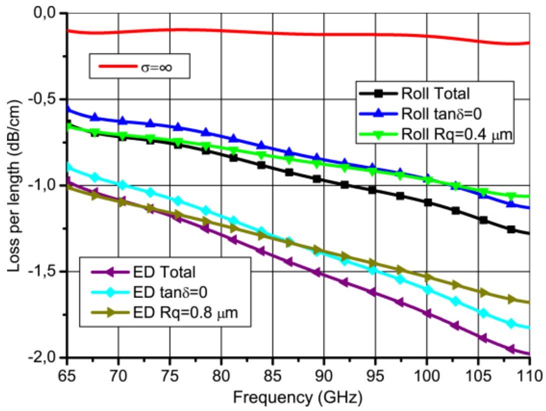


(b)

Sl. 5. (a) Merena faza kraćeg (crno) i dužeg voda (crveno) za slučaj rolovane bakarne folije, (b) merena faza kraćeg i dužeg voda za slučaj ED folije.

Merene fazne karakteristike vodova za različite tipove folija je data na Sl. 5, ΔL = 10 mm.

Procena doprinosa gubitaka od dielektrika je dobijena zamenom modela bakarne folije sa idealnim električnim provodnikom $\sigma = \infty$, Sl. 6. Dok se uticaj gubitaka u metalu može dobiti uzimanjem da je $\tan\delta = 0$. Po ovome modelu sa Sl. 6 se vidi da je uticaj gubitaka u dielektriku daleko manji nego u metalu. Vidi se da gradijentni model može dosta dobro da isprati promenu gubitaka sa frekvencijom. Očekivano, gubici su znatno veći za ED bakarnu foliju.



Sl. 6. Mereni totalni gubici za slučaj sa rolovano bakarnom folijom (Roll Total) i ED folijom (ED total), procena uticaja gubitaka u metalu i dielektriku za date folije, i poređenje sa odgovarajućim gradijentnim modelom.

III. Zaključak

U radu je prikazan je značaj merenja parametara materijala pre izrade samog prototipa antene. Deklarisani parametri materijala na 10 GHz nisu pouzdani za više učestanosti, naročito u pogledu promena karakteristika samog materijala na velikom opsegu učestanosti. Za tačan model potrebno je proceniti parametre materijala preko merenja u opsegu od interesa.

Zahvalnica

Autori se zahvaljuju Ministarstvu prosvete, nauke i tehnološkog razvoja Republike Srbije za finansiranje istraživanja.

Literatura

[1] D. M. Pozar, “*Microwave Engineering*”, 4th ed., Wiley, 2011.

[2] J. Coonrod and A. F. Horn III, “Understanding Dielectric Constant for Microwave PCB Materials,” *High Frequency Electronics*, July 2011.

[3] A. F. Horn III *et al*., “Effect of conductor profile on the insertion loss, phase constant, and dispersion in thin high frequency transmission lines”, *DesignCon 2010*, Santa Clara, February 2010.

[4] N. K. Das, S. M. Voda and D. M. Pozar, “Two Methods for the Measurement of Substrate Dielectric Constant,” in *IEEE Trans. Microw. Theory Techn.*, vol. 35, no. 7, pp. 636-642, Jul 1987.

[5] N. Boskovic *et al*., “Printed Frequency Scanning Antenna Arrays With Enhanced Frequency Sensitivity and Sidelobe Suppression,” *IEEE Trans. Antennas Propag*., vol. 65, no. 4, pp. 1757-1764, April 2017.

[6] Rogers Corporation, “Copper Foils for High Frequency Materials”, *Data Sheet*, 2015.

[7] F. Maradei, H. Ke and T. H. Hubing, “Full-Wave Model of Frequency-Dispersive Media With Debye Dispersion Relation by Circuit-Oriented FEM,” *IEEE Transactions on Electromagnetic Compatibility*, vol. 51, no. 2, pp. 312-319, May 2009.

[8] G. Gold and K. Helmreich, “A physical model for skin effect in rough surfaces,” *2012 42nd European Microwave Conference*, Amsterdam, 2012, pp. 1011-1014.

Abstract

Choice of the material play crucial role in the printed antenna design. This is especially critical on the W-band, where the loss mechanism is significantly enhanced, which is why a very small number of standard dielectrics can be used. To suppress parasite modes and surface waves, dielectric thickness must be very small relative to the frequency. Here, the effect of surface roughness of the copper conductor foil has a dominant impact on losses. This paper presents the importance of choosing an appropriate foil and following consequences. The conclusions are accompanied by simulations and measurements.

Material Characteristics for Printed Antenna Applications at 65-110 GHz

Nikola Boskovic and Milos Radovanovic

Influence of Various EM Models of an Aircraft to Monostatic RCS

Tomislav Milošević

***Abstract*—This paper outlines influence of four EM models of an electrically large aircraft on monostatic RCS results at 2.00 GHz. The differences in calculated RCS results suggest the importance of a choice of an EM model depending on the particular scope. The paper provides better understanding of some of EM scattering effects frequently addressed by engineering, scientific and military working groups interested in RCS. Software tool used for simulations and model manipulations is a full wave 3D EM Method-of-Moments based software with Surface Integral Equations applied to quadrilateral mesh elements.**

***Index Terms*—Scattering, aircraft, RCS, simulation.**

I. Introduction

CALCULATION of scattering from electrically large aircrafts is often a subject of interest of scientific, engineering, or military working groups. The purpose of scattering calculation can vary from an academic discussion, to simulation software testing, to anti-aircraft defense tactics preparation. Sometimes a calculation of scattering from electrically large fighter aircrafts is driven by marketing as various teams simulate fighter aircrafts in order to impress existing or future customers. According to information coming from various open sources, pure metallic models of aircrafts are still fashionable in various electromagnetic (EM) software tools (for example, the models of 4th generation of fighter jets are still widely used). They are usually modeled as metallic surfaces. Regarding the shape of an aircraft model, the details are usually classified and the shape usually comes rather from loose visual impression then from precise engineering data. For the sake of modeling some details are usually simplified, e.g., the engine intake is terminated with a metallic plate. The problem of inaccurate modeling arises when such model is chosen as the reference model in a realistic scenario.

The scattering results are usually related with radar operations and obtained after illuminating a target (an aircraft) with a plane wave. This paper will consider monostatic scattering and illumination with EM wave containing electric field with the E_θ component, only [1].

In monostatic radar setup the same antenna is used for both transmitting and receiving the signal. Complex radar targets such as aircrafts generally have (monostatic) cross sections that vary rapidly with frequency and aspect angle. Actually, a radar target is characterized by its radar cross section, which gives the ratio of scattered power to incident power density. The ratio depends on the target shape, the frequency and the polarization of the incident EM wave, and on the incident angle relative to the target. Monostatic radar cross section (RCS) can be defined as RCS where angles of incident and reflected waves are identical. In general, more complex targets require more efficient numerical techniques for software simulations [2].

Tomislav Milošević is with WIPL-D d.o.o., Gandijeva 7, 11073 Belgrade, Serbia (e-mail: tomislav.milosevic@wipl-d.com).

Here, we will consider four EM models of a single representative sample in the form of the 4th generation fighter aircraft. The EM models encompass scenarios where:

- The aircraft canopy and the radome are modeled as metallic surfaces while the engine intake is terminated with a metallic plate.
- The aircraft canopy and the radome are modeled as metallic surfaces while the engine intake can be considered as one-side open cavity with a metallic plate located in front of the jet engine rotor blades.
- The aircraft canopy is modeled as a metallic surface, the radome is excluded exposing a flat surface of the radar antenna while the engine intake is in the form of the open cavity.
- The aircraft canopy is excluded from the model exposing a pilot's working area which is also an open cavity. The radome is excluded exposing a flat surface of the radar antenna. The engine intake is in the form of the open cavity.

The results of monostatic RCS simulations of the four models will be compared and discussed. The aircraft dimensions and shape are to some degree approximate with the respect to actual aircraft.

The simulations will be facilitated in the frequency domain. In this paper, WIPL-D Software, a full wave, 3D EM frequency-domain Method-of-Moments (MoM) based software will be exploited for importing and modifying available CAD file and simulations applying higher order basis functions on quadrilateral mesh elements with Surface Integral Equations [3]. Since radar frequencies used for long range surveillance are located within L-band between 1 GHz and 2 GHz [4] and airport surveillance primary radar frequencies are about 2.8 GHz [5], the models presented here will be simulated in-between these frequencies trying to grasp EM effects appearing at both bands. Thus, the EM models of the aircraft are simulated at frequency of 2 GHz.

II. Aircraft EM Models

Four CAD models of the aircraft are shown in Figs. 1-4. All models are displayed with a symmetry plane. The symmetry plane represents a software feature where a symmetry of the structure is exploited to reduce an original number of unknowns. The original number of unknowns is approximately halved. Each figure also contains a magnified detail of an aircraft model. For all models it is assumed that aircraft surfaces are perfect electrically conductive metal. The models were imported and subsequently modified to have lower surfaces of wings smooth i.e., without pylons intended for carrying various loads.

The first model, which is shown in the Fig. 1 represents the model with the aircraft canopy and the radome modeled with metallic surfaces, while the engine intake is terminated with a metallic plate. This model is probably the most often seen in various software presentations and booklets. Fig. 1 also contains approximate dimension of the model which can be referenced when estimating the dimensions of the other models. In general, the model shown in the Fig. 1 is suitable for simulations with geometrical\physical optic-based EM solvers since the cavity in the form of engine intake is practically excluded from the simulation. From the marketing point of view this model mimics very well the realistic aircraft structure.

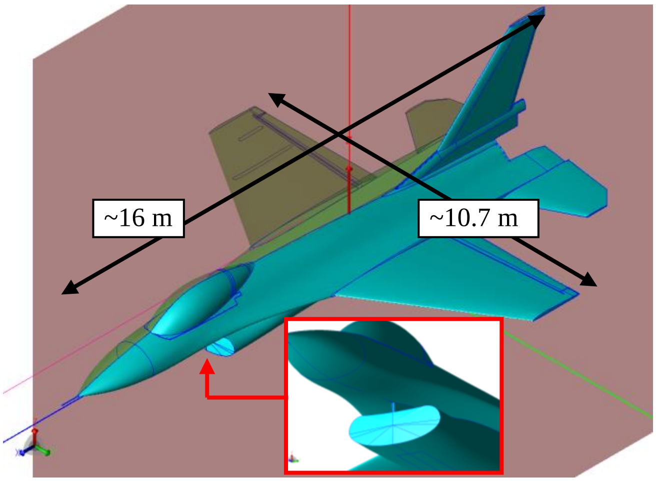


Fig. 1. The aircraft canopy and the radome are modeled as metallic surfaces while the engine intake is terminated with a metallic plate. This is probably the most commonly used EM model of an aircraft.

The second EM model shown in the Fig. 2 represents the model where the aircraft canopy and the radome are modeled as metallic surfaces while the engine intake can be considered as an open cavity with a metallic plate located in front of the jet engine rotor blades. Such model can also be often seen in various software presentations and booklets, despite presence of the cavity.

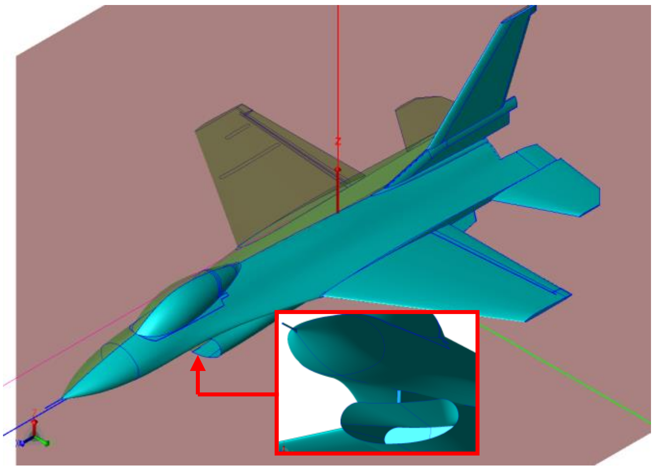


Fig. 2. The model of the aircraft which is also often simulated. The aircraft canopy and the radome are modeled as metallic surfaces while the engine intake can be considered as an open cavity with a metallic plate located in front of the jet engine rotor blades.

The third model is shown in the Fig. 3 and it can be used as a good representation for many types of aircrafts. In order to reduce influence of the pilot working area to RCS which also represents a sort of an open cavity (see also Fig. 4), some aircraft real-life models have canopy painted with the special material which is visually transparent and which increases radar waves reflection [6]. This model has the aircraft canopy modeled as a metallic surface, while the airborne radome covering radar antenna and providing aerodynamic streamlining is excluded from the model exposing radar antenna flat surface. Removing radome is justified as the radomes are generally composed of low-loss dielectrics materials [7]. The assumption applied here is that the radome is transparent for EM waves with frequency of 2 GHz. The engine intake is again in the form of an open cavity.

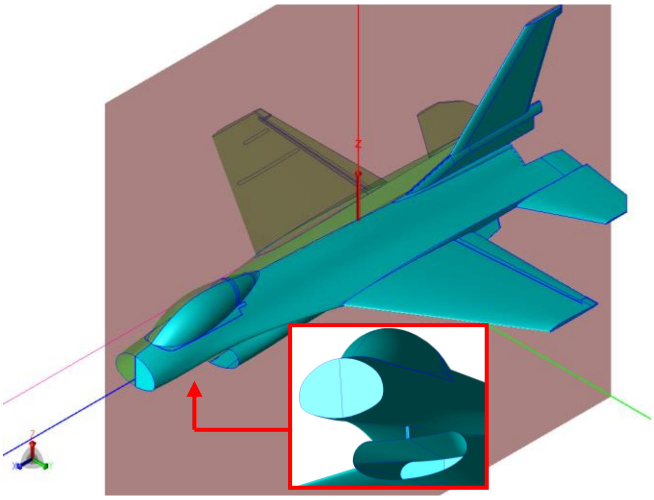


Fig. 3. The model of the aircraft where the aircraft canopy is modeled as a metallic surface, the radome is excluded exposing radar antenna flat surface while the engine intake can be considered as an open cavity with a metallic plate located in front of jet engine rotor blades.

Finally, the fourth EM model is shown in the Fig. 4.

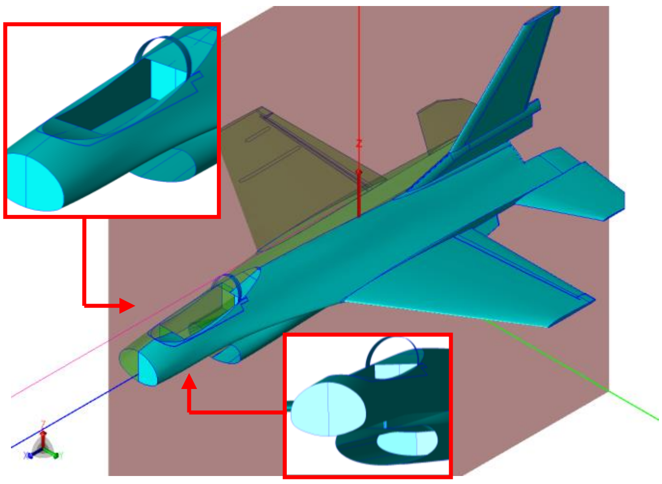


Fig. 4. The model of the aircraft which is probably the most rarely seen. The aircraft canopy is excluded from the model exposing an open cavity which represents a pilot’s working space. The radome is excluded exposing radar antenna flat surface. The engine intake is in the form of an open cavity with a metallic plate located in front of jet engine rotor blades.

The model shown in the Fig. 4 encompasses the aircraft canopy excluded from the model exposing cavity representing a pilot’s working area. The assumption applied here is that the canopy is transparent for EM waves with frequency of 2 GHz. Also, the radome is excluded exposing radar antenna flat surface. The engine intake is in the form of an open cavity.

In order to define suitable nomenclature of the models, four acronyms will be introduced. The model shown in the Fig. 1 will be named and referred further as MMM since the radome, the canopy, and the termination of the cavity are modeled with metallic surfaces (metal-metal-metal, respectively).

The model shown in the Fig. 2 will be named and referred further as MMA since the radome, and the canopy are modeled with metallic surfaces, while the termination of the cavity is excluded (it is assumed that it is modeled with air surface). In that sense, the name of the model will be MMA (metal-metal-air, respectively).

The model shown in the Fig. 3 will be named and referred further as AMA since the radome is replaced with air, the canopy is modeled with metallic surfaces while the termination of the cavity is excluded (again, it is assumed that it is modeled with air surface). In that sense, the name of the model will be AMA (air-metal-air, respectively).

The model shown in the Fig. 4 will be named and referred further as AAA since the radome, the aircraft canopy, and the termination of the cavity are all excluded from the model (assuming that all of them they are modeled with air surfaces). In that sense, the name of the model will be AAA (air-air-air, respectively).

Meshing details of the four models are presented in the Fig. 5-Fig. 8.

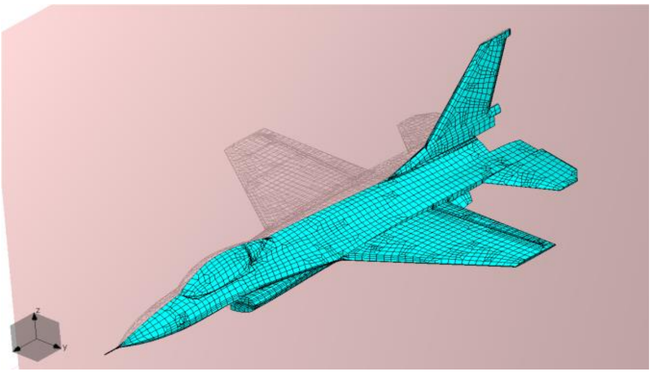


Fig. 5. Meshed metal-metal-metal (MMM) model

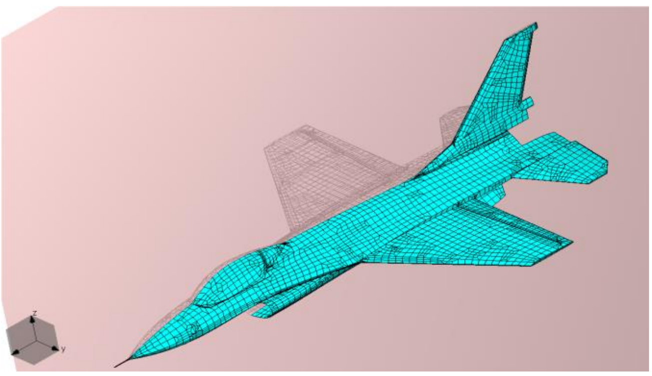


Fig. 6. Meshed metal-metal-air (MMA) model.

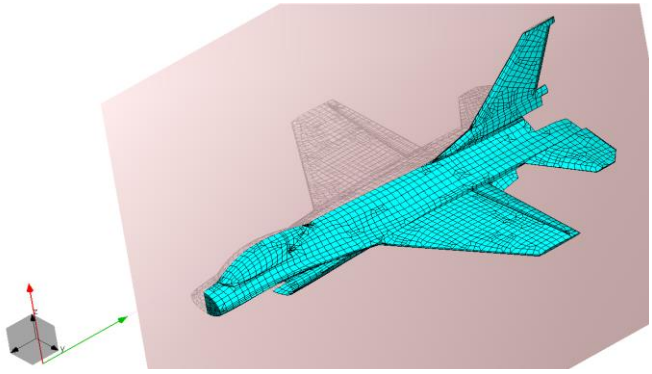


Fig. 7. Meshed air-metal-air (AMA) model.

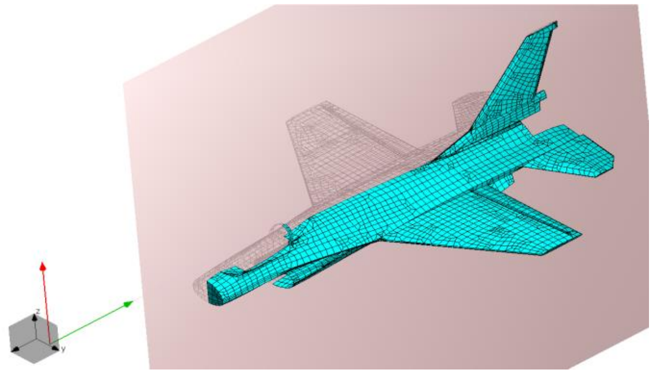


Fig. 8. Meshed air-air-air (AAA) model.

These four figures (Fig. 5-Fig. 8) depict meshed EM models i.e., models after applying mesh procedure and converting CAD files to simulation software native format. In the software native format, the aircrafts are modeled by using bilinear quadrilateral surfaces

III. Simulation Results

In order to discuss the influence of various aircraft modelling approaches, the calculated monostatic scattering results for four EM models follow. For the reasons of the careful comparison, it is convenient to present in the same graph the results originating from a pair of two models (three pairs in total, MMM-MMA, MMA-AMA and AMA-AAA). Eventually, MMM and AAA models will be compared.

All the results are obtained after calculating monostatic scattering from the front area of the aircraft. Actually, monostatic scattering is calculated in 901 directions encompassing theta angle span from 45 degrees below to 45 degrees above the aircraft nose. It is adopted that angle $\theta = 0$ degrees points toward aircraft nose (actually, it points toward horizon). The orientation of the aircraft and the theta angle are shown in Fig. 9.

All of the models have been simulated at operating frequency of 2 GHz. The workstation used for the simulations is Intel® Xeon® Gold 5118 CPU @ 2.30GHz 2.30 GHz (2 processors) with 192 GB RAM and 4 GPU cards Nvidia GeForce GTX 1080 Ti used for matrix inversion.

The most time-consuming simulation is MMA requiring less than 90 minutes to complete the simulation. Also, it requires 224,852 unknowns. The model MMM requires 218,601 unknowns while the model AMA requires 192,756 unknowns. Finally, the model AAA requires 193,419 unknowns.

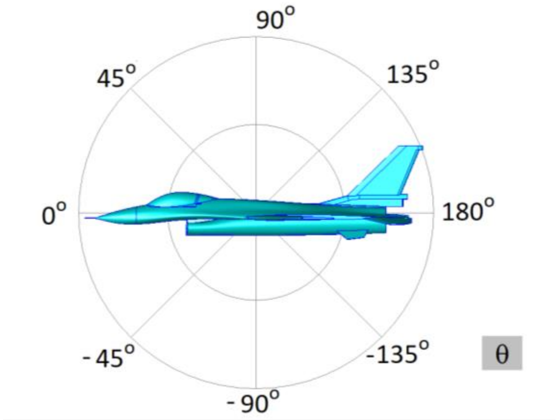


Fig. 9. Theta angle and the orientation of the aircraft.

A. *Comparing Results: MMM vs. MMA*

The comparison between MMM and MMA results is shown in the Fig. 10. A strong influence of cavity presence can be seen there. Actually, completely different results are obtained for angles between approximately -45 degrees and 5 degrees where, on average, the results differ by approximately 15 dB. This angle span corresponds to the angles of incidence important for ground-based surveillance radars.

The monostatic RCS for angles higher than about 10 degrees is almost the same for both models. It is expected since these directions are not affected by a way of terminating the engine intake, in this case located below the aircraft.

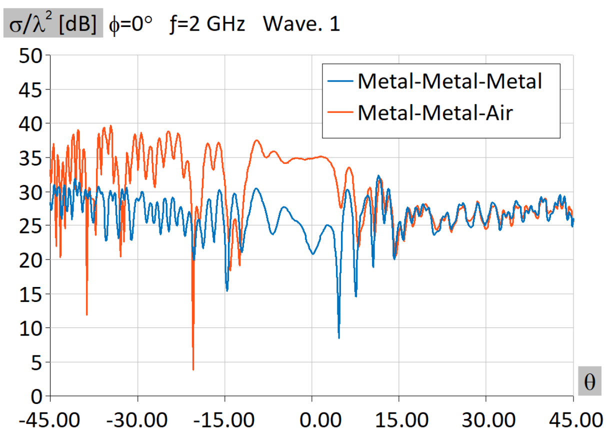


Fig. 10. Monostatic normalized RCS: MMM vs. MMA

B. *Comparing Results: MMA vs. AMA*

The comparison between MMA and AMA results is shown in Fig. 11. The only difference appears around 0 degrees. This is expected since the main difference between these two EM models appears in the area of the aircraft nose. This difference is significant in environment scenarios in which the aircraft is illuminated from the horizon (e.g., if an aircraft nose is illuminated from another airborne surveillance radar).

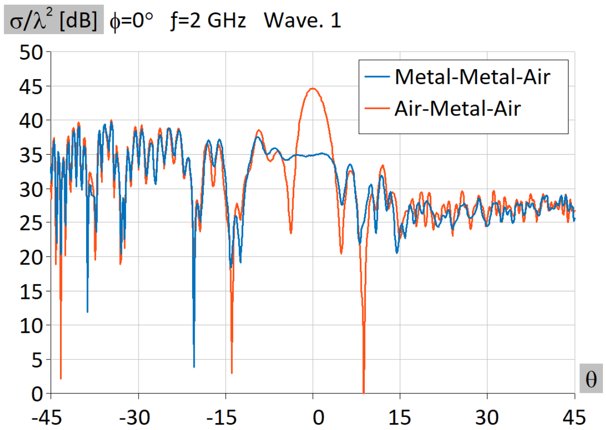


Fig. 11. Monostatic normalized RCS: MMA vs. AMA.

C. *Comparing Results: AMA vs. AAA*

The comparison between AMA and AAA results is shown in the Fig. 12. The influence of representing pilot’s working area as an open cavity is clearly seen in the scattering directions above theta angle of about 15 degrees. The reason for the differences can be explained similarly as in the MMM-MMA case - the presence of the cavity increases monostatic normalized RCS level.

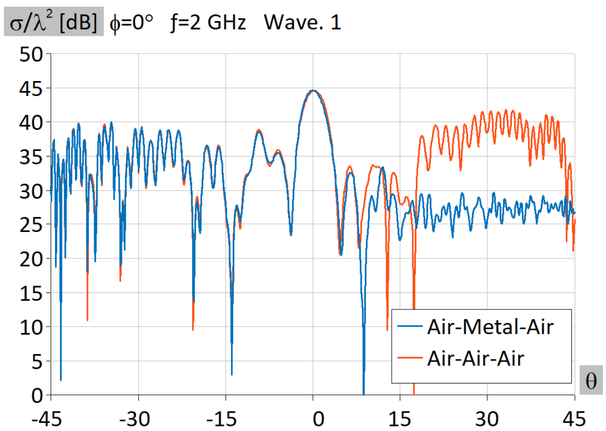


Fig. 12. Monostatic normalized RCS: AMA vs. AAA.

D. *Comparing Results: MMM vs. AAA*

In order to compare the two extreme cases considered in this paper, (MMM-AAA), calculated monostatic normalized RCS results are presented in the Fig. 13. The differences are considerable in the whole range of theta angles (-45 deg, 45 deg), but the origin for the differences can be easily tracked and explained in each of three mentioned subranges due to the previous three analysis and comparisons.

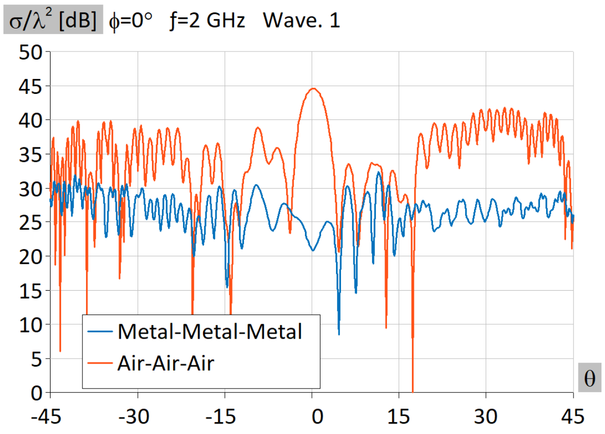


Fig. 13. Monostatic normalized RCS: MMM vs. AAA.

IV. Conclusion

This paper outlines influence of various EM models of an aircraft to monostatic normalized RCS results. The software tool used for the simulations and the model manipulations was a full wave 3D EM Method-of-Moments based software with Surface Integral Equations applied to quadrilateral mesh elements.

The simulated fighter aircraft is modeled using PEC metallic surfaces. The operating frequency is selected to be between frequencies of acquisition radar and frequencies of airport surveillance primary radar. The excitation is a linearly polarized EM plane wave containing only theta component of the electric field.

Four simulated EM models of the aircraft represent four common cases of aircraft models. Starting from a widely accepted model of the aircraft, named MMM here, the models are modified by adding or removing metallic surfaces, while dielectric properties of the surfaces are ignored.

A relevant EM model of the aircraft is related to a particular purpose, from the software marketing to the various military scenarios. From the marketing point of view, the adequate model contains the least modifications and it is easily understandable (MMM). However, for modeling real-life aircraft and obtaining results required for military purposes, the models similar to AMA or AAA should be used. The two models exhibit significant EM effects coming from the details such as jet engine intake, radome and aircraft canopy and are not expected at the first glance.

This aircraft represents to some extent a large and complex radar target. Thus, it can be assumed that in a specific angle span, monostatic RCS levels change if some parts of the aircraft are replaced with another parts. This assumption can be identified easily in the Fig. 11, where a flat plate produces significantly larger reflection compared to a conical shape representing metallic aircraft nose. An effect of this kind is expected, due to the nature of an EM wave scattering from the metallic surfaces of various shape. The similar effect can be noticed if a drop shaped canopy (or flat plate terminating the engine intake) is replaced with the open cavity. Since the aircraft is large and the reflections from these areas are assumed to be almost independent, all the effects presented in Fig. 13 using a single aircraft model can be also obtained by concatenating separately obtained results shown in Figs. 10-12.

The high efficiency of computation can be confirmed through the simulation times as they are all relatively short considering electrical size of the simulated structure and the workstation used.

The further investigation of this structure will include influence of the gun pipe and presence of various door openings to monostatic scattering. Also, further investigation will include application of radar-absorbing materials to selected aircraft surfaces in order to decrease a scattering level.

References

[1] B. M. Kolundzija, A. R. Djordjevic, *Electromagnetic Modeling of Composite Metallic and Dielectric Structures, 1st* ed. Norwood, Massachusetts, USA: Artech House, 2002.

[2] D. M. Pozar, *Microwave Engineering, 2nd* ed., USA, John Wiley & Sons, 1998.

[3] WIPL-D Software, WIPL-D d.o.o, Belgrade 2021. www.wipl-d.com

[4] https://www.radartutorial.eu/07.waves/Waves%20and%20Frequency%20Ranges.en.html

[5] https://en.wikipedia.org/wiki/Airport_surveillance_radar

[6] https://en.wikipedia.org/wiki/Aircraft_canopy

[7] J. L. Volakis, *Antenna Engineering Handbook, 4th* ed., USA: McGraw-Hill, 2007.

Utilization of Characteristic Mode Analysis in Coupled Resonators Microstrip Filter Design

Ana Đurđević and Milka Potrebić, *Member, IEEE*

***Abstract*—Examination of resonant frequency and coupling coefficient is essential part in microwave filters design with coupled resonators. We introduce Characteristic Mode Analysis (CMA) based method for calculation of coupling curve, applied to microstrip resonators. The main advantage of this approach is simplicity, due to CMs independency of any external sources. The results for coupling curve are presented and cross-checked with results obtained by equivalent two ports microstrip model with feeding lines. A very good agreement between the two methods is observed.**

***Index Terms*—Coupling coefficient, coupled resonators, CMA, filter, resonant frequency.**

I. Introduction

DISHAL acknowledged that any narrow-band bandpass filter can be described by tuning frequency of the resonators, the couplings between adjacent resonators and the external *Q*-factor of the first and last resonators [1], [2]. Initially, in case of distributed resonators, these parameters had usually been determined without a straight-forward method, but rather with a set of numerous experiments and measurements. With expansion of electromagnetic (EM) solvers, generating design curves for coupling and external *Q*-factor in terms of variables of interest became significantly easier. In [3], it was demonstrated how to extract the curves from insertion loss response and time delay at resonances, obtained from EM simulations. That approach we will later use in order to verify our results.

Since it was introduced, Characteristic Mode Analysis (CMA) was mainly used to analyze radiating properties of various antennas and scattering objects [4]–[6]. Apart from these fields, it found a role in analysis of surface wave resonance as well [7]. In this paper we outline a method for investigating resonant frequencies and generating coupling curve between resonators using CMA theory, in example of microstrip coupled resonators. We haven't yet considered a way to determine external *Q*-factor using CMA, which would allow to apply Dishal's concept [2] with all three variables required for bandpass filter design obtained by CMA only. The main purpose of this paper is to indicate a potential of CMA as a tool in narrow-band microwave bandpass filter design.

Chosen example for presenting the method consists of two grounded $\lambda/4$ resonators in microstrip technique, designed on operating frequency 1 GHz. To take advantage of Characteristic Modes formulation for PEC bodies [8], we considered air substrate. All the full-wave 3D EM analysis was performed in WIPL-D software package [9], which utilizes MoM (*Method of Moments*) and HOBFs (*Higher Order Basis Functions*).

Ana Djurdjevic is with WIPL-D d.o.o., Gandijeva 7, apt. 32, 11073 Belgrade, Serbia, and she is also a master student at the School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11020 Belgrade, Serbia (E-mail: ana.djurdjevic@wipl-d.com).

Milka Potrebic is with the School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11020 Belgrade, Serbia (E-mail: milka_potrebic@etf.rs).

II. Theoretical Insight to Characteristic Modes Analysis

Characteristic mode analysis is the numerical calculation of a weighted set of orthogonal current modes that are supported on a given structure. The theory was first introduced to electromagnetic by Garbacz [3], and later Harrington and Mautz formulated generalized eigenvalue equation for conducting bodies applied on MoM impedance matrix [8]. Here follows brief insight to derivation of eigenvalue equation. When an incident plane wave $\mathbf{E}^{\mathrm{i}}$ illuminates PEC structure, it induces surface currents $\mathbf{J}_{\mathrm{s}}$, which induce scattered field $\mathbf{E}^{\mathrm{s}}$. Boundary condition for $\mathbf{E}$ filed on the PEC body surface S can be written as:

$$(\mathbf{E}^{\mathrm{i}}(\mathbf{r})+\mathbf{E}^{\mathrm{s}}(\mathbf{r}))_{\tan}=0,\quad \mathbf{r}\in S, \tag{1}$$

where "tan" denotes tangential components of electric field. Scattering field $\mathbf{E}^{\mathrm{s}}$ can be expressed in terms of induced surface current as:

$$\begin{aligned}\mathbf{E}^{\mathrm{s}}=&-\frac{j\omega\mu_0}{4\pi}\int_S G(\mathbf{r},\mathbf{r}')\mathbf{J}_s(\mathbf{r}')dS'\\&-\frac{j}{4\pi\varepsilon_0\omega}\nabla\int_S G(\mathbf{r},\mathbf{r}')\nabla'\mathbf{J}_s(\mathbf{r}')dS'\end{aligned} \tag{2}$$

where ε_0 and μ_0 are the permittivity and permeability of the free space, and $G(\mathbf{r},\mathbf{r}')$ is Green's function in free space multiplied by 4π, and given by

$$G(\mathbf{r},\mathbf{r}')=\frac{e^{-jk_0R}}{R},\quad R=|\mathbf{R}|,\quad \mathbf{R}=\mathbf{r}-\mathbf{r}', \tag{3}$$

where R is distance between the field and the source point and $k_0=\sqrt{\varepsilon_0\mu_0}$ is the wavenumber in free space. Relationship between scattering field and surface currents can be written in form of integro-differential operator $L(\cdot)$, thus boundary condition (1) can be expressed as:

$$\left[L(\mathbf{J}_s)\right]_{\tan} = \mathbf{E}^{\mathrm{i}}_{\tan}(\mathbf{r}),\ \mathbf{r} \in S \tag{4}$$

that is known as EFIE (*Electric Field Integral Equation*). If we write tangential component of $L(\cdot)$ operator as a new operator $Z(\cdot)$, we obtain:

$$\left[L(\mathbf{J}_s)\right]_{\tan} = Z(\mathbf{J}_s). \tag{5}$$

Operator $Z(\cdot)$ has impedance property and can be split into real and imaginary part as:

$$\mathbf{Z} = \mathbf{R} + j\mathbf{X} \tag{6}$$

which represents MoM impedance matrix. By using impedance matrix in weighted eigenvalue equation, generalized eigenvalue equation for CM calculation is defined as:

$$\mathbf{X}(\mathbf{J}_{s,n}) = \lambda_n \mathbf{R}(\mathbf{J}_{s,n}). \tag{7}$$

Solutions of (7) are eigenvectors $\mathbf{J}_{s,n}$, that are vectors of current coefficients, eigenvalues λ_n, and n is the order of each mode. Eigenvalue is the real number within a range $\left[-\infty, +\infty\right]$, and its magnitude is proportional to the total stored field energy:

$$\omega \iiint_V (\mu \mathbf{H}_n \cdot \mathbf{H}_n^* - \varepsilon \mathbf{E}_n \cdot \mathbf{E}_n^*) \mathrm{d}V = \lambda_n. \tag{8}$$

Physical meaning of eigenvalues can be interpreted as follows:

- In the case of $\lambda_n = 0$, stored electric and magnetic energies are equal, and associated modes are considered as the resonant modes.
- In the case of $\lambda_n < 0$, stored electric energy dominates, and associated modes are considered as the capacitive modes.
- In the case of $\lambda_n > 0$, stored magnetic energy dominates, and associated modes are considered as the inductive modes.

The more convenient way for graphical representing eigenvalue is parameter called Modal Significance, defined as:

$$\mathrm{MS} = \left|\frac{1}{1 + j\lambda_n}\right| \tag{9}$$

MS takes values from [0,1], and for resonant modes, when λ_n approaches to 0, it is close to 1.

The important property which can be noticed from (7), is that CMs do not depend on any external excitation, but on the physical properties of the structure only. MoM matrix is filled-in, after which eigenvalue equation is solved in order to calculate unknown current coefficients for each mode, i.e. eigenvectors, as well as eigenvalues. Consequently, there is no need for any feeding network in analysis of the resonant properties in this way.

As it is mentioned before, the reason for analysis of PEC body in free space is to present the research using available tool [9]. In [10], theory of characteristic modes for material bodies is introduced, where the main difference from the theory for perfectly conducting bodies lies in the computation of the modes. It is also discussed in [10] that characteristic modes in material bodies have the most properties as corresponding modes in perfectly conducting bodies. Having that in mind, we may assume that the method is also valid for microstrip with other dielectric properties.

III. Generating Coupling Curve

Generally speaking, coupling coefficient can be defined as a ratio of coupled energy to stored energy, while corresponding electric and magnetic fields should be calculated at resonant frequencies [11]. In [12] it is given the expression for coupling coefficient, derived from lumped-element circuits:

$$k = \left(f_2^2 - f_1^2\right) / \left(f_1^2 + f_2^2\right), \tag{10}$$

where f_1 and f_2 are the lower and upper resonant frequency of the coupled resonators. Equation (10) is also valid for the distributed resonators, and resonant frequencies can be determined from full wave EM simulations. In following examples, coupling coefficient is always calculated using (10), while required resonant frequencies are obtained from both CMA and insertion loss response from model with feeding lines and ports. As an electromagnetic coupling increases when the elements are getting closer to each other, it is significant to graphically represent it as a function of distance, thus to generate coupling curve.

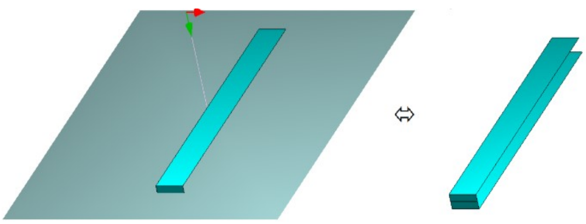


Fig. 1. Single grounded resonator above infinite PEC plane.

A. *CMA Based Method and Numerical Model*

Firstly, a single λ/4 resonator at operating frequency 1 GHz, with a grounded end above PEC plane, is analyzed with CMA solver. Distance from PEC plane is $h = 2$ mm, length of resonator $l = 74.95$ mm, and width $w = 9.83$ mm. In CMA model it is not possible to define a port between ground and resonators that would introduce a voltage necessary for transmission line. This obstacle is overcome by defining an infinite PEC plane that indicates the image theory to be applied. The original and the equivalent model after image theory are applied as given in Fig. 1.

Analysis is performed in discrete frequency points, in range from 0.8 GHz to 5 GHz. MS for the first 3 modes is shown in Fig. 2.

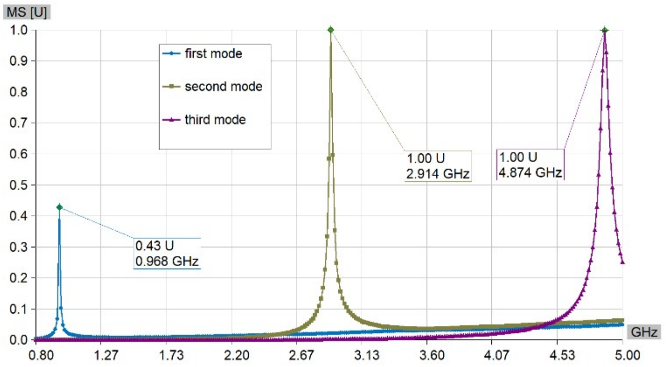


Fig. 2. MS for the first three modes of single resonator in frequency range 0.8 GHz – 5 GHz.

It can be seen from Fig. 2, that in the analyzed range there are three very narrow resonant modes, and the one of interest is at around 0.968 GHz. In order to analyze the coupling, one more resonator is added with the same dimensions, but grounded on different end, as shown in Fig. 3, and the results are given on Fig. 4. This model of coupled resonators is important for interdigital bandpass filter design.

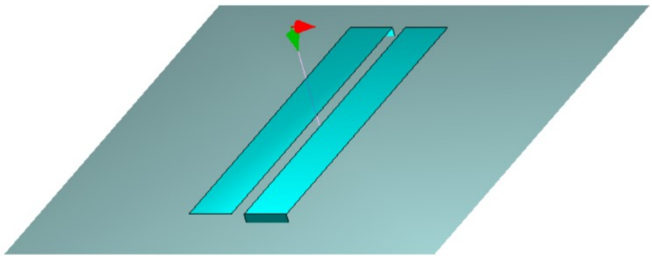


Fig. 3. Two coupled grounded resonators above infinite PEC plane.

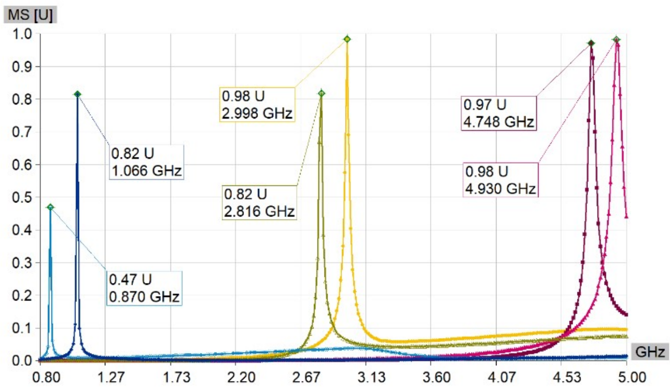


Fig. 4. MS for the first six modes of two coupled resonators in frequency range 0.8 GHz – 5 GHz.

As it was expected, instead of each resonance in case of single resonator, now we have two resonances shifted in frequency. Each peak still comes from different characteristic mode, but the two in every pair is very physically similar to each other, meaning have similar current and near field distribution. With varying the spacing between resonators, the spikes appear in the different positions in frequency. Having the low and high resonant frequencies as the only unknown quantities in (10), coupling coefficient can be calculated easily. In this example for spacing $s = 1$ mm, coupling coefficient equals to 0.2.

B. *Method Based on Insertion Loss Response*

In order to cross-check calculated results for coupling obtained by CMA method, we created the equivalent microstrip model with two resonators and two loosely coupled feed lines, presented in Fig. 5. The dielectric is air. The distances between feed lines and resonators equals to $2h$. Two resonators are grounded and two ports are placed on the different ends of feeding lines. The length and the width of the ground plane equal to 135 mm and 100 mm, respectively. The insertion loss response ($IL = -20\log_{10}|s_{21}|\mathrm{dB}$) is given in Fig. 6, and it can be observed that six resonances appeared at the same frequencies as resonances shown in MS response in Fig. 4.

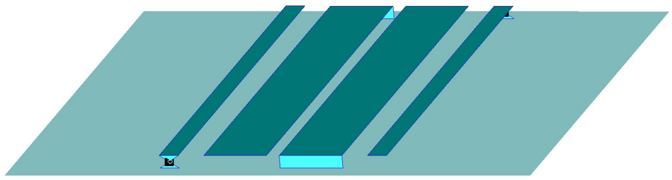


Fig. 5. Equivalent model with resonators and feeding lines.

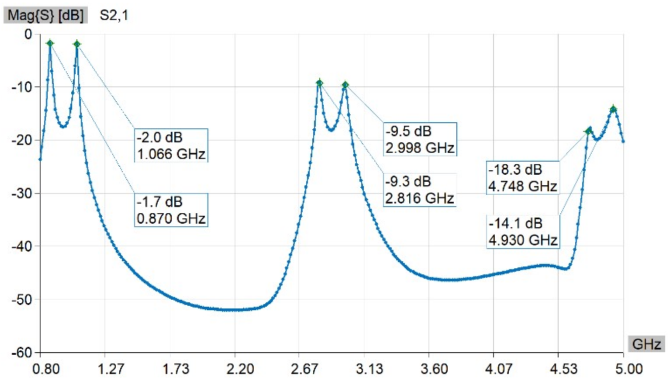


Fig. 6. Insertion loss response in frequency range 0.8 GHz – 5 GHz.

C. *Coupling Curve Results*

In case of both methods, models were re-simulated for different values of spacing between resonators, from range 0.1 mm to 10 mm. For each point, the equation (10) is calculated by inspecting resonances from MS graph and s-parameters. Results are overlaid on graph shown in Fig. 7, where “MS” denotes CMA based method, while “s_{21}” denotes the method based on insertion loss response.

The graph from Fig. 7 confirms that these two approaches result in the same coupling coefficient curves, and it can be said the CMA based method is verified. Negligible differences could be reduced by additional increasing of EM simulation accuracy, which is for the purpose of this research considered unnecessary.

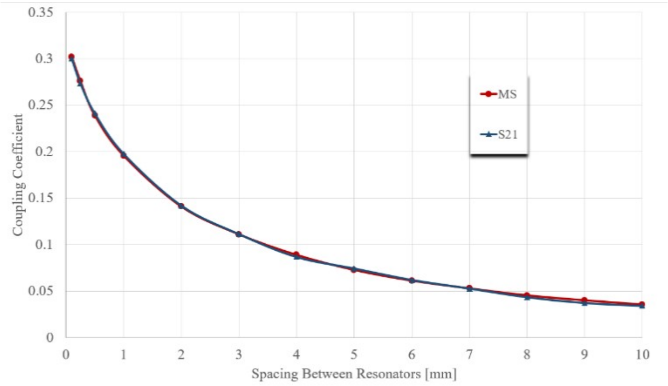


Fig. 7. Coupling curve in terms of spacing between resonators.

IV. Conclusion

In this paper, we have shown how CMA can be used in order to analyze resonant frequencies and coupling between resonators, and therefore can be used in bandpass microstrip filter synthesis. It provides very elegant solution to inspect internal resonances without taking care of feeding network and its potential influence. The validity of the results is confirmed with those obtained by inspecting *s*-parameters of the equivalent model with ports and feed lines.

Acknowledgment

This work was supported in part by the Ministry of Education, Science and Technological Development of the Republic of Serbia and by the Innovation Fund from the budget of the Republic of Serbia from the division of the Ministry of Education, Science and Technological Development, through the Serbia Competitiveness and Jobs Project (loan agreement with the World Bank).

References

[1] M. Dishal, “Alignment and Adjustment of Synchronously Tuned Multiple-Resonant-Circuit Filters,” *Proceedings of the IRE*, vol. 39, no. 11, pp. 1448-1455, Nov. 1951.

[2] M. Dishal, ”A Simple Design Procedure for Small Percentage Bandwidth Round-Rod Interdigital Filters (Correspondence),” *IEEE T. Microw. Theory*, vol. 13, no. 5, pp. 696-698, September 1965.

[3] R. J. Garbacz, R. H. Turpin, “A generalized expansion for radiated and scattered felds”, *IEEE T. Antenn. Propag.* vol. 19, no. 3, pp. 348-358, 1971.

[4] Q. Wu, “Characteristic Mode Assisted Design of Dielectric Resonator Antennas With Feedings,” *IEEE T. Antenn. Propag*, vol. 67, no. 8, pp. 5294-5304, 2019.

[5] J. J. Borchardt and T. C. Lapointe, “U-Slot Patch Antenna Principle and Design Methodology Using Characteristic Mode Analysis and Coupled Mode Theory,” *IEEE Access*, vol. 7, pp. 109375-109385, 2019.

[6] L. Guo, J. Zhao, C. Wang, Y. Chen and S. Yang, “Scattering Control Using Advanced Characteristic Mode Theories,” *2018 IEEE Int. Symp. Antenn. Propag. & USNC/URSI Nat. Radio Sci. Meeting*, 2018, pp. 1127-1128.

[7] Z. Xu, S. Li, Y. Liu, H. Zhao and X. Yin, “Characteristic Mode Analysis of Complex Spoof Localized Surface Plasmon Resonators,” *IEEE Access*, vol. 6, pp. 2871-2878, 2018.

[8] R. F. Harrington, J. R. Mautz, ”The Theory of Characteristic Modes for Conducting Bodies,” *IEEE T. Antenn. Propag.*, vol. 19, no. 5, pp. 622-628, 1971.

[9] WIPL-D Pro CAD 2020, WIPL-D d.o.o, Belgrade, 2020.

[10] R. Harrington, J. Mautz, Yu Chang, ”Characteristic modes for dielectric and magnetic bodies,” *IEEE T. Antenn. Propag.*, vol. 20, no. 2, pp. 194-198, March 1972.

[11] J.-S. Hong, “Couplings of asynchronously tuned coupled microwave resonators,” *IEE P. – Microw. Anten. P.*, vol. 147, no. 5, Oct. 2000, pp. 354–358.

[12] J.G. Hong, *Microstrip Filters for RF/Microwave Applications,* 2nd ed. Hoboken, New Jersey: Wiley, 2011.

АУТОМАТИКА
/
AUTOMATION

(АУ/AUI)

Ublažavanje četeringa digitalnog regulatora promenljive strukture za linearne sisteme

Boban Veselić, *Senior Member, IEEE*, Čedomir Milosavljević, Branislava Peruničić-Draženović, *Senior Member, IEEE*, Senad Huseinbegović, *Member, IEEE*, Milutin Petronijević, *Member, IEEE*

***Apstrakt*—U radu se ispituje mogućnost smanjenja četeringa u kliznom režimu jednog digitalnog regulatora promenljive strukture namenjenog za linearne sisteme sa ograničenim upravljačkim ulazom. Nelinearna diskontinualna funkcija relejnog tipa, koja je sastivni deo originalnog zakona upravljanja, zamenjena je adekvatnom kontinualnom aproksimacijom. Analizirana je stabilnost modifikovanog sistema i istaknute su uočene osobine sistema. Dobijeni teorijski zaključci su potvrđeni simulacionom proverom na numeričkom primeru.**

***Ključne reči*—Sistemi upravljanja promenljive strukture, diskretni klizni režimi, četering, kompenzacija poremećaja.**

I. Uvod

Sistemi upravljanja promenljive strukture (SUPS) [1] u radnom kliznom režimu (KR) [2] su jedna od popularnih robusnih nelinearnih tehnika upravljanja, teorijski invarijantni, u idealnom KR, na parametarske i spoljne poremećaje koji deluju u kanalu upravljanja [3]. Za ovu osobinu je potrebno diskontinualno upravljanje na kliznoj površi sa mogućnošću ostvarivanja beskonačne frekvencije preključivanja upravljačkih struktura. Iz tih razloga, zakoni upravljanja u KR sadrže diskontinualnu komponentu, najčešće relejnog tipa. U praktičnim realizacijama se osobina invarijantnosti redukuje u veliku robusnost sistema, upravo zbog realno ostvarljive visoke, ali konačne frekvencije preključivanja usled neidealnosti prekidačkih elemenata.

Glavna prepreka široke primene SUPS sa KR je pojava visokofrekvencijskih oscilacija u KR, poznatih kao četering, što je neprihvatljivo u nekim sistemima poput mehaničkih. Ovaj neželjeni efekat nastaje usled diskontinualne prirode upravljanja i postojanja neizbežne nemodelovane (parazitne) dinamike, koja se pobuđuje takvim upravljanjem i generiše četering.

U literaturi su predložene različite metode za ublažavanje četeringa u KR. Jedan od prvih pristupa [4] podrazumeva da se diskontinualna signum funkcija u upravljanju zameni kontinualnom nelinearnošću tipa zasićenja sa velikim pojačanjem, što za posledicu ima smanjenje robusnosti budući da pojačanje u sistemu postaje konačno. Drugi pristup je uvođenje opservera stanja [5], preko koga se zatvara petlja u kojoj ne egzistira četering. Petlja sa opserverom je idealna (nema nemodelovanu dinamiku) te nema četeringa u njoj pa služi kao bajpas za četering. Međutim, varijacije parametara sistema narušavaju robusnost i performanse sistema. Klizni režimi višeg reda (KRVR) [6], noviji pristup dosta popularan poslednju deceniju, su nastali u nastojanju rešavanja problema četeringa. Za nastanak KR r-tog reda potreno je da relativni red sistema u odnosu na kliznu promenljivu bude r, pa se tek u r-tom izvodu klizne promenljive javlja diskontinualno upravljanje. Naročitu pažnju je privukao tzv. super-twisting algoritam (STA) [6], razvijen za kontinualne sisteme sa jednim ulazom, koji ostvaruje KR drugog reda. Međutim, na osnovu analiza sprovedenim u [7-9] može se zaključiti da KRVR mogu ostvariti manji četering u odnosu na konvencionalne KR samo u sistemima sa dovoljno brzim aktuatorima. Dodatna redukcija četeringa se može ostvariti adekvatnim podešavanjem kontinualne aproksimacije signum funkcije u regulatorima promenljive strukture [10].

Digitalna implementacija upravljačkih algoritama uz pomoć mikroprocesora je otvorila analizu KR u vremenski diskretnom domenu [11]-[15]. Pokazano je da je samo u nominalnom sistemu moguće ostvariti diskretni KR, i pri tome upravljanje ne mora biti diskontinualno. U svim ostalim slučajevima moguće je ostvariti samo tzv. kvazi-KR (KKR), gde se kretanje sistema odvija u nekoj bliskoj okolini oko klizne površi. Kako vremenski diskretno upravljanje unosi dodatno kašnjenje u sistem, četering je izraženiji u slučaju diskretnih KR. Jedan od načina da se eliminiše ova pojava je da se u okolini klizne površi primeni linearno upravljanje [16], što će naravno dovesti do gubitka robusnosti sistema u toj zoni. U slučaju primene ove strategije, potrebno je izvršiti eventualnu kompenzaciju poremećaja u sistemu u cilju povećanja tačnosti. Za sporopromenljive poremećaje, kompenzaciona upravljanja su predložena u [17-19], koja imaju linearni karakter. U radu [20] je predložen digitalni regulator promenljive strukture sa nelinearnim kompezatorom poremećaja za linearne sisteme sa ograničem upravljačkim signalom. Predložena upravljačka struktura podseća na diskretizovanu varijantu STA upravljanja [21], ali indukuje manji četering.

U ovom radu se upravo polazi od digitalnog regulatora promenljive strukture [20], čije je upravljanje u okolini klizne

Boban Veselić i Milutin Petronijević – Elektronski fakultet, Univerzitet u Nišu, Niš, Srbija (e-mail: boban.veselic@elfak.ni.ac.rs, milutin.petronijevic@elfak.ni.ac.rs).

Čedomir Milosavljević – Elektrotehnički fakultet, Univerzitet u Istočnom Sarajevu, Istočno Sarajevo, Bosna i Hercegovina (e-mail: cedomir.milosavljevic@elfak.ni.ac.rs).

Branislava Peruničić-Draženović i Senad Huseinbegović – Elektrotehnički fakultet, Univerzitet u Sarajevu, Sarajevo, Bosna i Hercegovina (e-mails: brana_p@hotmail.com, shuseinbegovic@etf.unsa.ba).

Ovaj rad je podržan od strane Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije.

površi linearno i potpomognuto je kompenzacionim članom. Dopunska kompenzaciona upravljačka komponenta se formira nelinearnom estimacijom poremećaja, koja se dobija integracijom signuma klizne promenljive. Kako kompenzacioni deo regulatora ipak sadrži diskontinualnu funkciju, može doći do indukovanja četeringa. Iz tih razloga, u ovom radu se ispituje primena određene kontinualne aproksimacije signum funkcije sa ciljem ublažavanja četeringa, kao i njeni efekti na performanse sistema u pogledu stabilnosti i tačnosti. Određena je oblast konvergencije modifikovanog upravljačkog sistema i pokazano da se adekvatnim izborom parametra kontinualne aproksimacije može suziti širina kvazi-klizne oblasti u slučaju sporopromenljivih poremećaja u odnosu na originalni algoritam. Dobijeni teorijski rezultati su verifikovani simulacionim ispitivanjima.

II. OPIS DIGITALNOG REGULATORA PROMENLJIVE STRUKTURE

Posmatra se linearni kontinualni sistem upravljanja koji je opisan modelom u prostoru stanja

$$\dot{x}(t) = Ax(t) + b(u(t) + d(t)), \qquad (1)$$

gde je $x \in \mathbb{R}^{\boldsymbol{n}}$ vektor stanja, $u \in \mathbb{R}$ je upravljački signal, $A \in \mathbb{R}^{n\times n}$ i $b \in \mathbb{R}^{n\times 1}$ su matrice stanja i ulaza, respektivno. Na sistem deluje ograničeni poremećaj d, $|d(t)| \le d_0 < \infty$, koji zadovoljava uslove poklapanja, tj. deluje u prostoru upravljanja. Za realizaciju digitalnog regulatora potrebno je izvršiti vremensku diskretizaciju modela (1). Vremenski diskretni model sistema je oblika

$$x_{k+1} = A_d x_k + b_d(u_k + d_k), \qquad (2)$$

$$A_d = e^{AT},\ b_d = \int_0^T e^{At} B dt, \qquad (3)$$

pod pretpostavkom da je perioda diskretizacije T dovoljno mala i da su poremećaji u sistemu sporopromenljivi. Tada se poremećaj može smatrati konstantnim između dva uzastopna trenutka odabiranja, što za posledicu ima očuvanje uslova poklapanja u diskretnom domenu.

Da bi se ostvario KR željene dinamike potrebno je da se najpre dizajnira odgovarajuća klizna površ, a da se potom nađe upravljanje koje uspostavlja KR. Kretanje sistema se tada može dekomponovati na dve faze: fazu dosezanja klizne površi i fazu klizanja po datoj površi. Za svaku od ovih faza vezana je odgovarajuća upravljačka komponenta.

Pristup koji omogućava jasno razdvajanje komponenti upravljanja za dosezanje i klizanje bazira se na korišćenju diskretnog modela sistema koji je dobijen primenom δ-transformacije, [16]. Matematički model sistema (1) u δ-domenu se dobija korišćenjem modela (2) na sledeći način.

$$\delta x_k = \frac{x_{k+1} - x_k}{T} = A_\delta x_k + b_\delta(u_k + d_k), \qquad (4)$$

$$A_\delta = (A_d - I_n)/T,\ b_\delta = b_d/T. \qquad (5)$$

KR treba organizovati po površi $s_{\delta,k} = 0$ u prostoru stanja, gde se klizna promenljiva u δ-domenu definiše kao

$$s_{\delta,k} = c_\delta x_k,\ c_\delta b_\delta = 1. \qquad (6)$$

Željene sopstvene vrednosti sistema su definisane spektrom u kontinualnom domenu

$$\lambda = [\lambda_1 \quad \lambda_2 \quad \cdots \quad \lambda_{n-1} \quad 0], \qquad (7)$$

pri čemu nulta sopstvena vrednost ukazuje da je dinamika sistema u kliznom režimu redukovanog reda, tj. $n-1$ reda. Nenulte sopstvenene vrednosti se mogu preslikati u δ-domen na osnovu (5) kao $\lambda_{\delta i} = (e^{\lambda_i T} - 1)/T$, $i = 1, \cdots, n-1$, dok nulta sopstvena vrednost ostaje nepromenljiva. Spektar u δ-domenu je sada definisan sa

$$\lambda_\delta = [\lambda_{\delta,1} \quad \lambda_{\delta,2} \quad \cdots \quad \lambda_{\delta,n-1} \quad 0]. \qquad (8)$$

Formula za nalaženje vektora klizne površi [22], u ovom slučaju postaje

$$c_\delta = [k_{\delta e} \quad 1] \cdot [A_\delta \quad b_\delta]^\dagger, \qquad (9)$$

gde je $k_{\delta e}$ vector pojačanja povratne sprege po stanju koji obezbeđuje spektar (8) u sistemu (4), dok operator † označava pseudo-inverziju matrice.

Na osnovu (6) i (4) sledi da je

$$\delta s_{\delta,k} = \frac{s_{\delta,k+1} - s_{\delta,k}}{T} = c_\delta \delta x_k = c_\delta A_\delta x_k + u_k + d_k. \qquad (10)$$

Ekvivalentno upravljanje u δ-domenu se određuje rešavanjem uslova $s_{\delta,k+1} = 0$ po upravljanju u_k, te se korišćenjem prethodne jednačine dobija

$$u_{eq,k} = -c_\delta A_\delta x_k - \frac{s_{\delta,k}}{T} - d_k. \qquad (11)$$

Član $s_{\delta,k}/T$ u upravljanju je zadužen za fazu dosezanja, i postaje jednak nuli kada sistem dođe na kliznu površ. Ovom komponentom može da se utiče na fazu dosezanja.

Ostvarljiv deo ekvivalentnog upravljanja (11) je

$$u_k = -c_\delta A_\delta x_k - \frac{s_{\delta,k}}{T} = -(k_{\delta e} + \frac{1}{T} c_\delta) x_k. \qquad (12)$$

Zadatak upravljanja je obezbediti brz odziv bez preskoka i značajnu robusnost na poremećaje u sistemu. Takođe treba uzeti u obzir neizbežno postojanje zasićenja u aktuatorima, što uslovljava ograničavanje upravljačkog signala regulatora po amplitudi. Navedeni zahtevi su ostvareni digitalnim regulatorom promenljive strukture sa kliznim režimom [20]. Upravljačka strategija ovog regulatora se sastoji iz dva režima rada. U prvoj fazi, kada je stanje sistema udaljeno od klizne površi, deluje linearno upravljanje sa zadatkom da dovede stanje sistema u blisku okolinu klizne površi koja zavisi od amplitude poremećaja koji deluje. Nakon toga se aktivira nelinearna kompenzacija poremećaja koja ojačava linearnu upravljačku komponentu. Na taj način se izbegava preskok i pojava integratorskog zamaha usled postojanja nelinearnosti tipa zasićenja.

Zakon upravljanja [20] je opisan sledećim jednačinama:

$$u_k = \begin{cases} U_0 \operatorname{sgn}(u_{\Sigma,k}), & |u_{\Sigma,k}| > U_0 \\ u_{\Sigma,k}, & |u_{\Sigma,k}| \le U_0 \end{cases}, \qquad (13a)$$

$$u_{\Sigma,k} = u_{l,k} - p_{2,k} u_{c,k}, \qquad (13b)$$

$$u_{l,k} = -c_\delta A_\delta x_k - T^{-1}[k_{s1} + (1 - p_{2,k}) k_{s2}] s_{\delta,k}, \qquad (13c)$$

$$u_{c,k} = u_{c,k-1} + k_{int} T \operatorname{sgn}(s_{\delta,k-1}) \qquad (13d)$$

$$p_{1,k} = \begin{cases} 0, & |u_{\Sigma,k}| > U_0 \\ 1, & |u_{\Sigma,k}| \le U_0 \end{cases}, \qquad (13e)$$

$$p_{2,k} = p_{1,k-1}, \qquad (13f)$$

$$k_{s1}, k_{s2} > 0,\ k_{s1} + k_{s2} \le 1. \qquad (13g)$$

Može se primetiti da se ukupno upravljanje $u_{\Sigma,k}$ sastoji iz dve komponente: linearne $u_{l,k}$ i nelinearne (kompenzacione) $u_{c,k}$. Kada je stanje sistema daleko od klizne površi, izlaz regulatora u_k je u zasićenju jer linearno upravljanje teži da dovede sistem na kliznu površ u nominalnom slučaju. Izračunato upravljanje $u_{\Sigma,k}$ je velike amplitude koja se onda limitira na U_0, vrednost koja je prihvatljiva za aktuator. Nakon izlaska regulatora iz zasićenja primenjuje se linearno upravljanje

$$u_{l,k} = -c_\delta A_\delta x_k - T^{-1}(k_{s1} + k_{s2})s_{\delta,k} \quad (14)$$

u trajanju od jedne periode diskretizacije. Ovaj mehanizam je ostvaren pomoćnim promenljivama $p_{1,k}$ $p_{2,k}$. U graničkom slučaju kada su pojačanja $k_{s1} + k_{s2} = 1$, upravljanje (14) postaje tzv. dead-beat upravljanje koje daje $s_{\delta,k+1} = 0$ u slučaju nominalnog sistema $(d_k = 0)$. U realnom sistemu uzima se da je $k_{s1} + k_{s2} \leq 1$ pošto je $d_k \neq 0$ (uključujući i nemodelovanu dinamiku) da bi se podesila širina kvazi-kliznog domena. Ovakvo upravljanje dovodi stanje sistema u neku okolinu klizne površi. U sledećoj periodi diskretizacije redukuje se pojačanje linearnog dela upravljanja i aktivira se nelinearna komponenta, tako da se tada upravljanje definiše realacijama

$$u_k = -c_\delta A_\delta x_k - T^{-1}k_{s1}s_{\delta,k} - u_{c,k}, \quad (15)$$

$$u_{c,k} = u_{c,k-1} + k_{int}T\mathrm{sgn}(s_{\delta,k-1}). \quad (16)$$

Komponenta upravljanja $u_{c,k}$ je izlaz diskretnog integratora sa pojačanjem k_{int} koja teži da kompenzuje dejstvo poremećaja. Dakle, (16) sprovodi nelinearnu estimaciju poremećaja. Na ulaz integratora se dovodi signum funkcija klizne promenljive.

Stabilnost sistema podrazumeva konvergenciju trajektorija sistema ka kliznoj površi u oba režima rada regulatora, u zasićenju i van zasićenja. U [20] su izvedeni uslovi stabilnosti sistema, dati kroz dva tvrđenja. Prvo tvrđenje se odnosi na potrebnu veličinu limita izlaza regulatora.

Tvrđenje 1: Diskretni sistem (4) sa regulatorom (13) koji radi u režimu zasićenja će napustiti ovaj režim za konačan broj perioda uzorkovanja ukoliko važi da je

$$U_0 > |c_\delta A_\delta x_k| + d_0, \forall k \geq 0. \quad (17)$$

Drugo tvrđenje sagledava uslove konvergencije nakon izlaska regulatora iz režima zasićenja.

Tvrđenje 2: Uslovi konvergencije diskretnog sistema (4) sa regulatorom (13), koji radi van režima zasićenja i ima parametre $0 < k_{s1} \leq 1$ i $k_{int} > 0$, će biti ispunjeni unutar oblasti definisane sa

$$|s_{\delta,k}| > \frac{k_{int}T^2}{k_{s1}}. \quad (18)$$

Iz ovog uslova se vidi da za konačno malo T postoji uska oblast oko klizne površi u kojoj ne postoji konvergencija trajektorija ka kliznoj površi i koja utiče na širinu kvazi-klizne oblasti. Može se zaključiti da je prilikom projektovanja sistema poželjno ostvariti što je moguće manje T, jer ne postoji bojazan generisanje prevelikog upravljanja i potencijalne nestabilnosti usled postojanja limitera na izlazu regulatora.

III. Ublažavanje četeringa

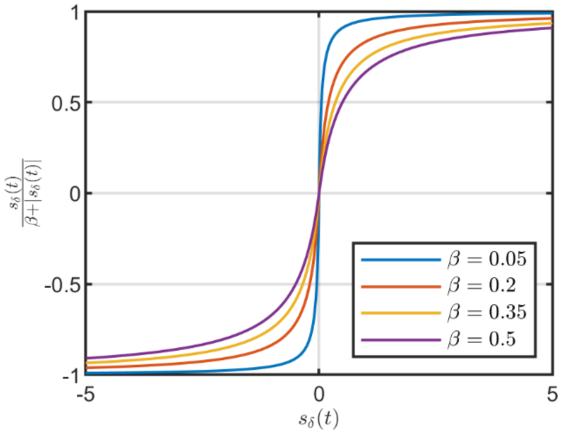
Relacije (15) i (16) opisuju rad regulatora (13) u željenom režimu rada van zasićenja. U estimatoru poremećaja (16) figuriše diskontinualna funkcija. Iako diskontinualni signal $\mathrm{sgn}(s_{\delta,k-1})$ prolazi kroz integrator, diskretna realizacija integratora ne može u potpunosti da isfiltrira ovaj signal. Zato ovakav estimator ipak predstavlja izvor četeringa u sistemu.

U ovom radu se ispituje mogućnost dalje redukcije četeringa zamenom diskontinualne signum funkcije u (13d), odnosno (16), nekom kontinualnom aproksimacijom. Kako se signum funkcija može opisati relacijom $\mathrm{sgn}(s_{\delta,k}) = s_{\delta,k}/|s_{\delta,k}|$, ideja je da se primeni u (16) sledeća kontinualna aproksimacija

$$u_{c,k} = u_{c,k-1} + k_{int}T\frac{s_{\delta,k-1}}{\beta + |s_{\delta,k-1}|}, \beta \geq 0. \quad (19)$$

Uticaj parametra β na oblik aproksimacije je prikazan na Sl. 1. Očigledno je da za $\beta = 0$, (19) se svodi na (16), tj. nema aproksimacije.

Nakon uvođenja kontinualne aproksimacije, potrebno je analizirati stabilnost i performanse sistema. Najpre se posmatra rad regulatora (13) sa aproksimacijom (19) u režimu zasićenja. Kako tada aproksimacija (19) nije aktivna i upravljanje je maksimalne amplitude $u_k = U_0\mathrm{sgn}(u_{\Sigma,k})$, uvođenje aproksimacije nema uticaj na ovaj režim, te i dalje važe uslovi (17). Dalje treba analizirati rad sistema van zasićenja upravljačkog signala.



Sl. 1. Izgled kontinualne aproksimacija funkcije $\mathrm{sgn}(s_{\delta,k})$ za različite vrednosti parametra $\beta > 0$.

Nakon izlaska upravljanja iz zasićenja, na osnovu zakona (13) u narednoj periodi diskretizaije deluje upravljanje

$$u_k = -c_\delta A_\delta x_k - T^{-1}(k_{s1} + k_{s2})s_{\delta,k}, \quad (20)$$

pa $\delta s_{\delta,k}$ postaje

$$\delta s_{\delta,k} = -T^{-1}(k_{s1} + k_{s2})s_{\delta,k} + d_k. \quad (21)$$

Pošto je $\delta s_{\delta,k} = T^{-1}(s_{\delta,k+1} - s_{\delta,k})$, iz prethodne jednačine se može naći $s_{\delta,k+1}$ kao

$$s_{\delta,k+1} = (1 - k_{s1} - k_{s2})s_{\delta,k} + Td_k. \quad (22)$$

Važno je primetiti da se u graničnom slučaju $(k_{s1} + k_{s2} = 1)$ dobija $s_{\delta,k+1} = Td_k$, što ukazuje da ovo upravljanje dovodi sistem u $O(T)$ okolinu klizne površi, čije dimenzije zavise i od amplitude poremećaja. Ukoliko nema poremećaja $(d_k = 0)$, dobija se $s_{\delta,k+1} = 0$ te je ovo upravljanje ustvari ekvivalentno upravljanje koje ostvaruje dead-beat odziv.

Već u narednoj periodi diskretizacije, pojačanje regulatora se redukuje $(k_{s2} = 0)$ i aktivira se kompenzaciono upravljanje $u_{c,k}$ (19). Izlaz regulatora se sada formira na osnovu jednačina

$$u_k = -c_\delta A_\delta x_k - T^{-1}k_{s1}s_{\delta,k} - u_{c,k}, \quad (23)$$

$$u_{c,k} = u_{c,k-1} + k_{int}T\frac{s_{\delta,k-1}}{\beta + |s_{\delta,k-1}|}, \quad (24)$$

pa se dinamika klizne promenljive opisuje sa

$$s_{\delta,k+1} = (1 - k_{s1})s_{\delta,k} + T(d_k - u_{c,k}), \quad (25)$$

$$u_{c,k+1} = u_{c,k} + k_{int}T\frac{s_{\delta,k-1}}{\beta + |s_{\delta,k-1}|}. \quad (26)$$

Kompenzaciono upravljanje $u_{c,k}$ ustvari predstavlja estimaciju poremećaja. Ako se uvede greška estimacije kao $z_k = d_k - u_{c,k}$, dinamika sistema (25), (26) se može izraziti kao

$$s_{\delta,k+1} = (1-k_{s1})s_{\delta,k} + Tz_k, \tag{27}$$

$$z_{k+1} = -k_{int}T\frac{s_{\delta,k-1}}{\beta+|s_{\delta,k-1}|} + z_k + \Delta_k, \tag{28}$$

gde je $\Delta_k = d_{k+1} - d_k$.

Stabilnost nelinearnog sistema (27), (28) se može analizirati korišćenjem pseudo-linearne forme [21], [23], što u ovom slučaju daje sledeći model

$$\sigma_{k+1} = \Lambda(s_{\delta,k})\sigma_k + p_k, \tag{29}$$

$$\sigma_k = \begin{bmatrix} s_{\delta,k} \\ z_k \end{bmatrix},\ \Lambda(s_{\delta,k}) = \begin{bmatrix} 1-k_{s1} & T \\ \frac{-k_{int}T}{\beta+|s_{\delta,k}|} & 1 \end{bmatrix},\ p_k = \begin{bmatrix} 0 \\ \Delta_k \end{bmatrix}$$

Karakteristična jednačina pseudo-linearnog sistema (29) se nalazi iz uslova $F(z) = \det(zI - \Lambda) = 0$, iz koga se dobija

$$F(z) = a_2z^2 + a_1z + a_0(s_{\delta,k}) = 0, a_2 > 0, \tag{30}$$

$$a_2 = 1, a_1 = k_{s1} - 2, a_0 = 1 - k_{s1} + \frac{k_{int}T^2}{\beta+|s_{\delta,k}|}.$$

Uslovi stabilnosti sistema se mogu odrediti primenom Džurijevog testa stabilnosti, koji se za sistem drugog reda (29) svodi na zahteve $F(1) > 0$, $F(-1) > 0$ i $|a_0| < a_2$. Za karakterističnog polinoma (30), ovi uslovi su sledeće relacije

$$\frac{k_{int}T^2}{\beta+|s_{\delta,k}|} > 0, \tag{31}$$

$$2(2-k_{s1}) + \frac{k_{int}T^2}{\beta+|s_{\delta,k}|} > 0, \tag{32}$$

$$\left|1 - k_{s1} + \frac{k_{int}T^2}{\beta+|s_{\delta,k}|}\right| < 1. \tag{33}$$

Kako važi da je $k_{int} > 0$, $k_{s1} \le 1$, $\beta \ge 0$ i $T > 0$, prve dve nejednakosti su uvek ispunjene. Dakle, uslov stabilnosti definiše (33). Pošto je izraz unutar apsolutne vrednosti u (33) uvek pozitivan, apsolutna vrednost se može zanemariti i uslov postaje $k_{int}T^2/(\beta + |s_{\delta,k}|) < k_{s1}$, iz koga se nalazi

$$|s_{\delta,k}| > \frac{k_{int}T^2}{k_{s1}} - \beta. \tag{34}$$

Dakle, sistem (29) je stabilan unutar oblasti definisane relacijom (34), što znači da će u toj oblasti trajektorije sistema biti usmerene ka kliznoj površi $s_\delta = 0$. Važno je istaći da u slučaju konstantnih ili sporopromenljivih poremećaja, za koje se može smatrati da važi $d_k = d_{k-1}$, sistem (29) je autonoman jer je $\Delta_k = 0$. Tada će trajektorije sistema dosegnuti granicu konvergencije $|s_{\delta,k}| = k_{int}T^2/k_{s1} - \beta$. Očigledno je da u neposrednoj okolini klizne površi, datoj sa $|s_{\delta,k}| < k_{int}T^2/k_{s1} - \beta$, uslovi konvergencije nisu zadovoljeni, pa trajektorije napuštaju ovu oblast i tako nastaje kvazi-klizno kretanje.

Iz (34) se vidi da se primenom aproksimacije okarakterisane sa β, pored redukcije četeringa, može povećati oblast konvergencije sistema, tj. smanjiti zona oko klizne površi gde se konvergencija ne garantuje. Interesantno je da se za vrednost $\beta = k_{int}T^2/k_{s1}$ dobija uslov konvergencije

$$|s_{\delta,k}| > 0, \tag{35}$$

koji ukazuje da su u celom prostoru oko klizne površi ispunjeni uslovi konvergencije. Tada se javlja diskretni KR. U slučaju $\beta = 0$, oblast konvergencije (34) se svodi na (18), što ukazuje na valjanost analize.

Ovde treba istaći da izvedeni zaključci važe u slučaju konstantnih i sporopromenljivih poremećaja. Za ostale tipove poremećaja treba očekivati da će njihova amplituda i frekvencija uticati na širinu kvazi-klizne oblasti, koja će svakako biti veća nego teorijski dobijena širina na osnovu oblasti konvergencije (34).

IV. Numerički primer i simulacioni rezultati

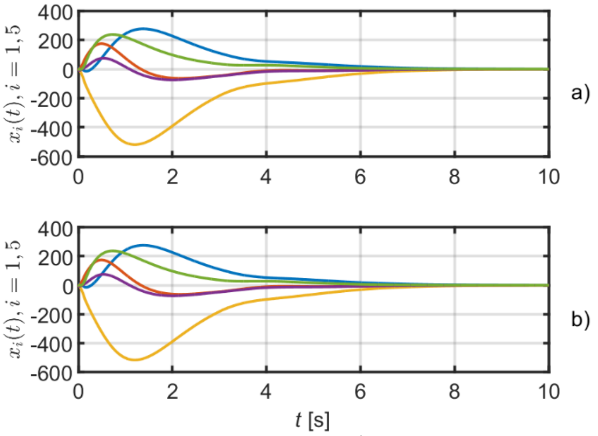
Dobijeni teorijski rezultati su provereni digitalnom simulacijom. Za tu svrhu je izabran akademski primer linearnog, nestabilnog i kontrolabilnog objetka upravljanja petog reda, čiji model (1) je definisan matricama

$$A = \begin{bmatrix} 1 & 2 & 3 & -5 & 6 \\ -2 & 6 & -3 & -4 & -7 \\ 2 & -4 & 6 & -10 & 12 \\ -8 & -6 & -4 & 3 & 1 \\ 4 & 12 & -6 & -8 & -14 \end{bmatrix}, b = \begin{bmatrix} 1 \\ -2 \\ 3 \\ -1 \\ 2 \end{bmatrix}.$$

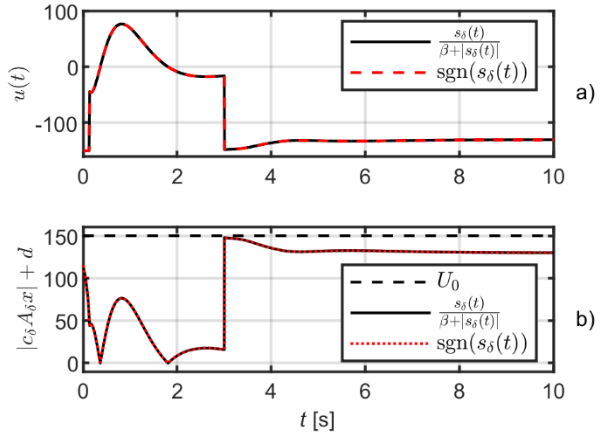
Neka je proizvoljno izabrano početno stanje sistema $x(0) = [10 \quad -5 \quad 0 \quad -10 \quad 5]^{\mathrm{T}}$. Perioda diskretizacije sistema je $T = 0.001$ s.

Zadatak upravljanja je da se sistema iz proizvoljnog početnog stanja dovede u ranotežno stanje (koordinatni početak), organizovanjem KR duž površi sa definisanom dinamikom. Neka je željena dinamika sistema u KR data spektrom sopstvenih vrednosti $\lambda = [-1 \quad -2 \quad -3 \quad -4 \quad 0]$ u kontinualnom domenu. Korišćenjem $\lambda_{\delta i} = (e^{\lambda_i T} - 1)/T$, $i = 1, \cdots, n-1$, dobija se spektar u δ-domenu

$$\lambda_\delta = [-0.9995 \quad -1.998 \quad -2.9955 \quad -3.992 \quad 0].$$



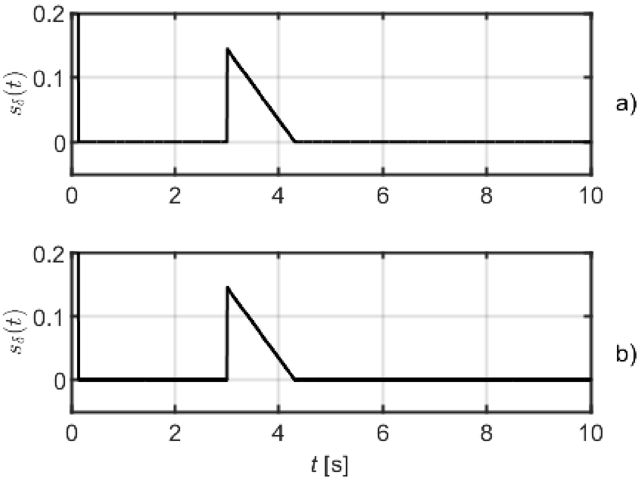
Sl. 2. Koordinate stanja: a) $\beta = 1.1111 \cdot 10^{-4}$; b) $\beta = 0$



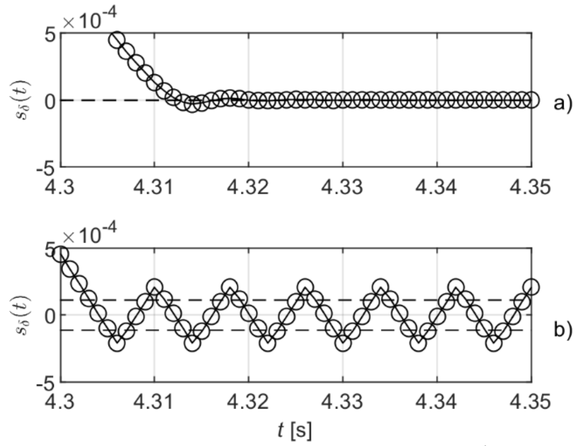
Sl. 3. a) Upravljanja; b) Ispunjenje uslova (17)

Odgovarajuća klizna površ koja obezbeđuje željenu dinamiku je definisana vektorom

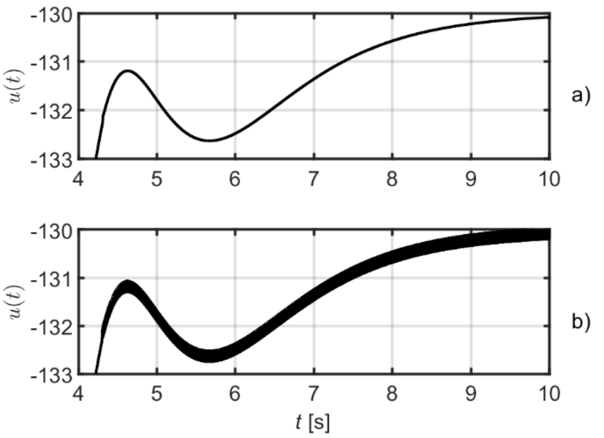
$c_\delta = [0.4437 \quad 0.5794 \quad 0.3072 \quad -0.6614 \quad 0.063]$, dobijenog iz (9). Parametri regulatora su $k_{s1} = 0.9$, $k_{s2} = 0.1$ i $k_{int} = 100$, dok je ograničenje u sistemu $U_0 = 150$.



Sl. 4 Klizna promenljiva: a) $\beta = 1.1111 \cdot 10^{-4}$; b) $\beta = 0$



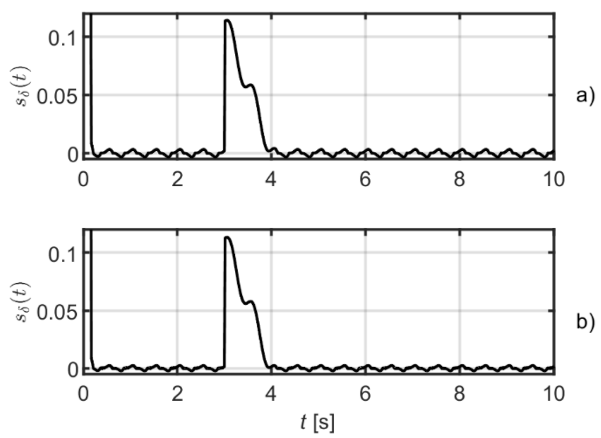
Sl. 5 Klizna promenljiva (uvećan detalj): a) $\beta = 1.1111 \cdot 10^{-4}$; b) $\beta = 0$



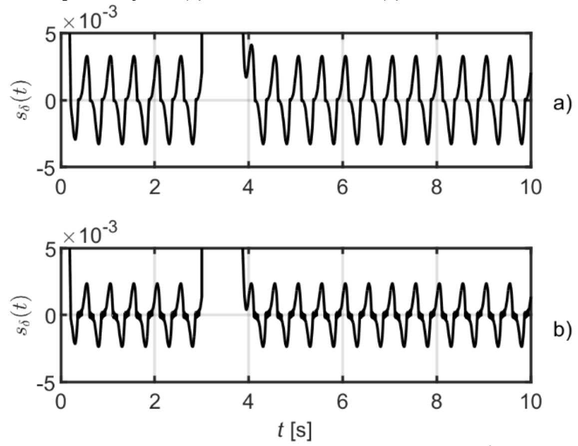
Sl. 6. Upravljanje (uvećan detalj): $\beta = 1.1111 \cdot 10^{-4}$; b) $\beta = 0$

U prvom testu sistem je izložen odskočnom poremećaju $d(t) = 130h(t-3)$, gde je $h(t)$ Hevisajdova funkcija. Posmatrani su slučajevi primene aproksimacije (19) sa karakterističnim vrednošću $\beta = k_{int}T^2/k_{s1} = 1.1111 \cdot 10^{-4}$ i slučaj sa signum funkcijom ($\beta = 0$). Na Sl. 2 i Sl. 3 su prikazane koordinate stanja, upravljanja i ispunjenje uslova izlaska sistema iz zasićenja (17), na kojima oba sistema imaju naizgled iste odzive. Koordinate stanja asimptotski teže koordinatnom početku uprkos dejstvu poremećaja. Izlazi regulatora su inicijalno u zasićenju ali veoma brzo izlaze iz zasićenja, budući da je uslov (17) sve vreme ispunjen. Klizne promenljive, prikazane na Sl. 4, ukazuju da se brzo uspostavlja KR i da dejstvo poremećaja izbacuje sistem iz KR ali da regulatori uspevaju da ponovo uspostave KR.

Međutim, uvećani detalji kliznih promenljivih na Sl. 5, gde su kružićima označene vrednosti klizne promenljive u trenucima odabiranja, ukazuju na različitu prirodu kretanja sistema u neposrednoj okolini klizne površi. Naime, za slučaj $\beta = k_{int}T^2/k_{s1}$ oblast konvergencije (35) obuhvata čitavu okolinu klizne površi, te trajektorija sistema treba da dođe na kliznu površ za dati tip poremećaja. Ovo predviđeno ponašanje sistema je potvrđeno odzivom na Sl. 5a, gde se u sistemu uspostavlja diskretni KR, pa je ostvarena maksimalna tačnost. U slučaju primene signum funkcije u regulatoru ($\beta = 0$), na osnovu (18) se vidi da u naznačenoj okolini klizne površi, oivičenoj isprekidanim linijama na Sl. 5b, ne postoji konvergencija te trajektorije sistema napuštaju ovu oblast. U ovom slučaju uspostavlja se diskretni kvazi-klizni režime, što ukazuje na smanjenu tačnost sistema i pojavu četeringa.



Sl. 7. Klizna promenljiva: a) $\beta = 1.1111 \cdot 10^{-4}$; b) $\beta = 0$



Sl. 8. Klizna promenljiva (uvećan detalj): a) $\beta = 1.1111 \cdot 10^{-4}$; b) $\beta = 0$

Efekti uvođenja kontinualne aproksimacije na redukciju četeringa se vide na Sl. 6, gde su prikazani uvećani detalji upravljačkih signala oba sistema. Očigledno je da je upravljanje u slučaju primene aproksimacije za graničnu vrednost β u potpunosti glatko, dok za $\beta = 0$ upravljanje sadrži visokofrekvencijsku komponentu po uspostavljanju kvazi-kliznog režima. Ova komponenta, iako je po amplitudi veoma mala, ipak ukazuje na mogućnost pojave četeringa.

U drugom testu se ispituju performanse sistema u prisustvu promenljivih poremećaja. Na sisteme deluje složeni poremećaj koji se sastoji iz konstantnog i sinusoidalnog dela, tj. $d(t) = 10\sin(4\pi t) + 100h(t-3)$. Ostala podešenja u sistemima su identična prethodnom slučaju.

Odzivi kliznih promenljivih su dati na Sl. 7 iz kojih se vidi da sistemi ispoljavaju manju robusnost na promenljive poremećaje u odnosu na sporopromenljive poremećaje. Na osnovu uvećanih delova ovih grafika, prikazanih na Sl. 8, vidi se da u oba slučaja nastaje kvazi-klizni režim. Takođe se vidi da je širina kvaziklizne oblasti u slučaju primene aproksimacije nešto veća u odnosu na slučaj za $\beta = 0$, iako za graničnu vrednost β su uslovi konvergencije ostvareni u celom prostoru. To ukazuje da za ovakav tip poremećaja sistem sa kontinualnom aproksimacijom zbog glatke kompenzacione komponente upravljana gubi izvesni stepen robusnosti, a samim tim i tačnosti.

V. Zaključak

Sprovedenom analizom i simulacionim rezultatima je pokazano da primena predložene kontinualne aproksimacije signum funkcije u digitalnom regulatoru promenljive strukture, pored efekta ublažavanja četeringa, otvara mogućnost proširivanja oblasti konvergencije sistema ka kliznoj površi. U slučaju sporopromenljivih poremećaja data aproksimacija ima blagotvorno dejstvo, jer se redukuje četering i ostvaruje diskretni KR, čime se postiže apsolutna tačnost. U slučaju promenljivih poremećaja, primena aproksimacije koja redukuje četering ipak dovodi do izvesnog smanjenja robusnosti, odnosno tačnosti. U tom slučaju je potrebno je naći kompromis između prihvatljivog nivoa četeringa i zahtevane tačnosti.

Literatura

[1] S. V. Emelyanov, A Method to Obtain Complex Regulation Laws Using Only the Error Signal or Regulated Coordinate and Its First Derivatives, Avtomat. Telemekh. 18 (10), pp. 873-885, 1957

[2] V.I. Utkin, Sliding Modes in Control and Optimization, Springer-Verlag, 1992.

[3] B. Draženović, (1969) The invariance conditions in variable structure systems. Automatica 5, pp. 287–295

[4] Slotine J.J.E. (1984) Sliding controller design for non-linear systems. Int. Jour. Control 40, pp. 421–434.

[5] Bondarev, A.G., Bondarev, S.A., Kostylyeva, N.Y., Utkin, V.I., “Sliding modes in systems with asymptotic observers”, Aut. Remote Contr., 46 (1985) no 6, pp. 679-684

[6] Levant A (1993) Sliding order and sliding accuracy in sliding mode control. Int. Jour. Control, 58, pp. 1247–1263.

[7] Utkin V., "Discussion aspects of higher order sliding mode control", *IEEE Trans.* AC, vol. 61, no. 3, pp. 829-833, 2016.

[8] Perez-Ventura U., Fridman L., “When is it reasonable to implement the discontinuous sliding-mode controllers instead of the continuous ones? Frequency domain criteria,” *Inter. Jour. of Robust and Nonlinear Control*, vol. 29, pp. 810-828, 2019.

[9] Utkin V., Poznyak A, Orlov Y., Polyakov A., “Conventional and high order sliding mode control,” *Jour. of the Franklin Institute*, vol. 357, no. 15, pp. 10244-10261, 2020.

[10] Castillo I., Freidovich L.B., “Describing-function-based analysis to tune parameters of chattering reduction approximations of sliding mode controllers,” *Control Engineering Practice*, vol. 95, 104230, 2020.

[11] Milosavljević, Č. (1985): "General Conditions for the Existence of Quasi-sliding Mode on the Switching Hyper-plane in Discrete Variable Structure Systems", *Automatic and Remote Control*, 46, pp. 307-314

[12] Drakunov, S. V., Utkin, V.I. (1989), ‘On discrete-time sliding mode’, Proc. IFAC Symposium on *Nonlinear Control Systems design*, Capry (Italy), pp. 484-489

[13] Gao, W., Wang, Y. Homaifa A. (1995), “Discrete-time variable structure control systems,*” IEEE Trans. Ind. Electron.*, IE-42, pp. 117-122.

[14] Bartolini G, Ferrara A, Utkin VI (1995) Adaptive sliding mode control in discrete-time systems. Automatica, 31, pp. 769–773.

[15] A. Bartoszewitz, “Discrete-time quasi-sliding mode control strategies,” IEEE Trans. Ind. Elect., vol. 45, pp. 633-637, 1998.

[16] Golo G, Milosavljević Č (2000) Robust discrete-time chattering free sliding mode control. Syst. Control Lett. 41, pp. 19–28.

[17] Su W.C., Drakunov S V., Özgüner Ü (2000) An O(T2) boundary layer in sliding mode for sampled-data systems. IEEE Trans. Automat. Contr. 45, pp.4 82–485.

[18] Milosavljević Č, Peruničić-Draženović B, Veselić B, Mitić D (2007) A new design of servomechanisms with digital sliding mode. Electr. Eng. 89, pp. 233–244.

[19] Lješnjanin M, Peruničić B, Milosavljević Č, Veselić B, (2011), Disturbance compensation in digital sliding mode. Int. conf. EUROCON. Lisboa, Portugal (paper 171)

[20] B. Veselić, Č. Milosavljević, B. Peruničić-Draženović, A. Huseinbegović, M. Petronijević, “Discrete-time sliding mode control of linear systems with input saturation,” *Int. J. Appl. Math. Comput. Sci*., vol. 30, no. 5, pp. 517-528, 2020.

[21] Koch S, Reichhartinger M, Horn M, Discrete-time equivalents of the super-twisting algorithm, Automatica 107 (2019) 190–199

[22] Draženović B, Milosavljević Č, Veselić B (2013) Comprehensive Approach to Sliding Mode Design and Analysis in Linear Systems. In: Bandyopadhyay B, Janardhanan S, Spurgeon SK (eds) Advances in Sliding Mode Control: Concept, Theory and Implementation. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 1–19

[23] Ghane H, Menhaj M.B, Eigenstructure-based analysis for non-linear autonomous systems, *IMA Journal of Mathematical Control and Information* (2015) 32, 21–40.

Abstract

The paper explores possibility of chattering reduction of a digital sliding mode controller for linear systems with saturated inputs. A nonlinear discontinuous relay function in the control law is substituted by an adequate continuous approximation. Stability of the modified control system is analyzed and the observed system properties are emphasized. The obtained theoretical results have been confirmed by simulation of a numerical example.

Chattering Alleviation of a Digital Sliding Mode Controller for Linear Systems

B. Veselić, Č. Milosavljević, B. Peruničić-Draženović, S. Huseinbegović, M. Petronijević

Autonomno kretanje besposadnog vozila po zadatoj putanji primenom algoritma sa aktivnim potiskivanjem poremećaja

Momir Stanković i Stojadin Manojlović

***Apstrakt*—U radu je predložen algoritam autonomnog kretanja besposadnog vozila po zadatoj putanji primenom koncepta upravljanja sa aktivnim potiskivanjem poremećaja. Nelinearnosti kinematike kretanja i poremećaji linearne i ugaone brzine vozila su formulisani u vidu totalnog poremećaja. Za estimaciju stanja nominalnog modela i totalnog poremećaja, kao dodatne promenljive stanja, projektovan je prošireni opserver stanja. Na bazi estimacija i dostupnih merenja formulisan je upravljački zakon za aktivno potiskivanje totalnog poremećaja u realnom vremenu i upravljanje kretanjem vozila sa definisanom dinamikom praćenja zadate putanje. Simulacionom analizom predloženog algoritma sa tipičnim modelom pogona guseničnog besposadnog vozila je pokazana efikasnost predloženog rešenja u različitim scenarijima praćenja zadate putanje.**

***Ključne reči*—Autonomno kretanje; Besposadno vozilo; Upravljanje sa aktivnim potiskivanjem poremećaja; Prošireni opserver stanja;.**

I. Uvod

Iako su ideje o automatizovanim pokretnim platformama bez ljudske posade (eng. *Unmanned Vehicles - UVs*) postojale još početkom prošlog veka, tek njegovim krajem došlo se do tehnoloških mogućnosti za njihovu realizaciju. Osnovna namena besposadnih platformi je delimična ili potpuna zamena čoveka u napornim, rizičnim, nepristupačnim i sl. civilnim ili vojnim misijama. Uz tehnološki napredak u oblastima senzorskih i aktuatorskih komponenti, obrade signala i algoritama vođenja (navigacije) i upravljanja, zbog potencijalnih mogućnosti ovih specifičnih platformi, poslednjih godina se poklanja sve veća pažnja njihovom razvoju i primeni. Posebnu grupu automatizovanih platformi predstavljaju besposadna vozila (eng. *Unmanned Ground Vehicles - UGVs*) projektovana za kretanje po uređenom ili neuređenom terenu. Razvoj besposadnih vozila u početku je bio zasnovan na automatizaciji kretanja postojećih vozila kojima je upravljao čovek. Međutim, zbog povećanja ekonomičnosti, mobilnosti i manevrabilnosti, danas se besposadna vozila projektuju kao posebna klasa vozila sa točkovima ili gusenicama, sa potpuno automatizovanim funkcijama kretanja i, najčešće, nezavisnim upravljanjem pogonskim točkovima [1].

Momir Stanković – Vojna akademija, Univerzitet odbrane u Beogradu, Generala Pavla Jurišića Šturma 33, 11000 Beograd, Srbija (e-mail: momir_stankovic@yahoo.com).

Stojadin Manojlović – Vojna akademija, Univerzitet odbrane u Beogradu, Generala Pavla Jurišića Šturma 33, 11000 Beograd, Srbija (e-mail: colemanojle@yahoo.com).

Sistem za autonomno kretanje besposadnog vozila se, u opštem slučaju, može razdvojiti na tri celine: sistem za analizu terena, sistem za planiranje (proračun) trajektorije i sistem upravljanja kretanjem. Sistem upravljanja kretanjem treba da obezbedi praćenje zadate trajektorije uz minimalna odstupanja. Funkcionalno se može razdvojiti na upravljanje uzdužnim kretanjem, odnosno upravljanje intenzitetom brzine kretanja i na upravljanje ugaonim kretanjem (uglom skretanja) [2]. Algoritam upravljanja treba da obezbedi robusnost i stabilnost kretanja, uzimajući u obzir dinamičko ponašanje i konstruktivna ograničenja samog vozila kao i uticaje poremećaja. Rešenja ovog problema variraju od primene klasičnih do inteligentnih tehnika upravljanja [3].

U ovom radu analiziran je algoritam upravljanja ugaonim kretanjem guseničnog besposadnog vozila za praćenje zadate putanje, primenom koncepta sa aktivnim potiskivanjem poremećaja (eng. *Active Disturbance Rejection Control - ADRC*). Ovaj koncept se pokazao kao veoma robusan, efikasan i praktičan u potiskivanju kako spoljašnjih (ambijentalnih), tako i unutrašnjih (sistemskih) poremećaja [4]. Pri modelovanju sistema, sva odstupanja od nominalnog modela, uključujući greške modelovanja (neodređenost/promene parametara, nemodelovana dinamika) i spoljašnje poremećaje, tretiraju se kao *totalni poremećaj sistema*. Potiskivanjem totalnog poremećaja preko unutrašnje povratne sprege, za kompenzovani (nominalni) sistem, u vidu redne veze integratora, projektuje se kontroler zatvaranjem povratnih sprega po stanjima. Klučna komponenta ADRC kontrolera je prošireni opserver stanja (eng. *Extended State Observer - ESO*), koji se projektuje za nominalni model, a totalni poremećaj se uvodi kao dodatno stanje sistema. Od kvaliteta estimacija ESO-a u najvećoj meri zavisi efikasnost ADRC kontrolera [5].

U radu je projektovan ADRC kontroler na osnovu kinematičkog modela kretanja besposadnog vozila, uz pretpostavke da je intenzitet brzine kretanja konstantan i da se praćenje referentne putanje ostvaruje kontrolisanjem ugaone orijentacije vozila. Pri tome se svi poremećaji koji mogu nastati u toku kretanja vozila usled proklizavanja, neravnog terena ili neadekvatnog odziva pogonskih motora, modeluju kao poremećaji intenziteta brzine vozila i poremećaji ugaone brzine njegove orijentacije. S obzirom da poremećaji intenziteta brzine vozila ne deluju na istom ulazu gde i upravljački signal (eng. *mismatched uncertainty*) model sistema je formulisan u pogodnom obliku kako bi se svi poremećaji uključili u totalni poremećaj.

Za estimaciju stanja nominalnog modela i totalnog poremećaja projektovan je linearni ESO. Koristeći estimacije i dostupne merene veličine, formulisan je zakon upravljanja za minimizaciju normalne komponente greške praćenja (eng. *cross-track error*), pri čemu se besposadno vozilo približava zadatoj putanji zahtevanom dinamikom.

II. Kinematički model kretanja vozila po zadatoj putanji

Kinematički model autonomnog kretanja vozila u inercijalnom koordinatnom sistemu se može opisati jednačinama:

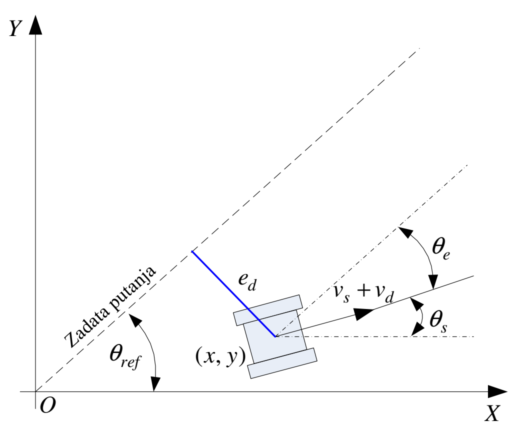
$$\begin{aligned} \dot{x}(t) &= \left(v_s(t)+v_d(t)\right)\cos\theta_s(t) \\ \dot{y}(t) &= \left(v_s(t)+v_d(t)\right)\sin\theta_s(t) \\ \dot{\theta}_s(t) &= \omega_s(t)+\omega_d(t) \end{aligned} \tag{1}$$

gde su $x(t)$ i $y(t)$ koordinate a $\theta_s(t)$ ugaona orijentacija vozila. Linearna brzina vozila v_s se smatra poznatom uz pretpostavku da je konstantna, a ugaona brzina vozila $\omega_s(t)$ predstavlja upravljačku veličinu, koju treba proračunati tako da vozilo izvrši zahtevani manevar. Uticaj proklizavanja, neravnog terena ili neadekvatnog odziva pogonskih motora na kretanje vozila modelovani su poremećajima linearne $v_d(t)$ i ugaone brzine $\omega_d(t)$. Poremećaji $v_d(t)$ i $\omega_d(t)$ predstavljaju osnovni uzrok odstupanja vozila od zadate putanje pri autonomnom kretanju vozila.

Na Sl. 1 prikazana je zadata putanja kao pravac definisan uglom θ_{ref} u inercijalnom koordinatnom sistemu. U tom slučaju je greška ugaone orijentacije vozila u odnosu na zadatu putanju:

$$\theta_e(t) = \theta_s(t) - \theta_{ref} \tag{2}$$

Poziciono odstupanje vozila definisano je greškom $e_d(t)$ kao normalno rastojanje vozila od zadate putanje.



Sl. 1. Zadata putanja kretanja vozila i greške praćenja.

Problem praćenja zadate trajektorije se može definisati kao regulacioni problem minimizacije greške $e_d(t)$ u prisustvu poremećaja $v_d(t)$ i $\omega_d(t)$ primenom odgovarajućeg upravljačkog signala $\omega_s(t)$. Na osnovu Sl. 1 i primenom izvoda na jednačinu (2), dobijaju se izrazi za dinamiku grešaka pozicionog odstupanja i ugaone orijentacije vozila u odnosu na zadatu putanju:

$$\dot{e}_d(t) = \left(v_s + v_d(t)\right)\sin\theta_e(t) \tag{3}$$

$$\dot{\theta}_e(t) = \omega_s(t) + \omega_d(t) \tag{4}$$

Imajući u vidu da je ugaona brzina rotacije vozila $\omega_s(t)$ upravljačka veličina, iz (3) se može uočiti da poremećaj $v_d(t)$ ne deluje na istom ulazu kao i $\omega_s(t)$ (eng. *mismatched uncertainty*), za razliku od poremećaja $\omega_d(t)$ (eng. *matched uncertainty*) [6].

Diferenciranjem (3), i uvrštavanjem (4), dobija se model dinamike greške $e_d(t)$:

$$\begin{aligned} \ddot{e}_d(t) &= \omega_s(t)(v_s+v_d(t))\cos(\theta_e) \\ &+\omega_d(t)(v_s+v_d(t))\cos(\theta_e)+\dot{v}_d(t)\sin(\theta_e) \end{aligned} \tag{5}$$

u kome je uticaj oba spoljašnja poremećaja sveden na zajednički ulaz sa upravljačkim signalom $\omega_s(t)$ (eng. *matched uncertainty*).

Iz (5) se vidi, da je dinamika greške $e_d(t)$, čak i u odsustvu poremećaja ($v_d(t) \equiv \omega_d(t) \equiv 0$), nelinearna. Ako se pretpostavi postojanje početne greške pozicioniranja $e_d(t_0) \neq 0$, željena dinamika minimizacije greške se može definisati kao:

$$\ddot{e}_d^*(t) + k_2\dot{e}_d^*(t) + k_1 e_d^*(t) = 0 \tag{6}$$

gde se pomoću k_1 i k_2 definišu parametri prelaznog procesa. Prema tome, zadatak praćenja zadate putanje predstavlja projektovanje upravljačkog signala $\omega_s(t)$ kojim se obezbeđuje da greška $e_d(t)$ sa zadovoljavajućom tačnošću prati zadatu dinamiku, definisanu sa (6).

III. Projektovanje ADRC regulatora za autonomno kretanje vozila

U skladu sa konceptom ADRC-a, izraz (5) se može zapisati u formi:

$$\ddot{e}_d(t) = v_s\cos\left(\theta_e(t)\right)\cdot\omega_s(t) + f(t) \tag{7}$$

gde je:

$$f(t)=\cos(\theta_e)v_d(t)\omega_s(t)+\omega_d(t)(v_s+v_d(t))\cos(\theta_e)$$
$$+\dot{v}_d(t)\sin(\theta_e)\ , \qquad (8)$$

totalni pormećaj, koji obuhvata nelinearnost modela i pretpostavljene spoljašnje poremećaje. Ako se $f(t)$ usvoji kao dodatno stanje sistema, (7) se može predstaviti modelom u prostoru stanja:

$$\dot{\mathbf{x}}(t)=\mathbf{A}\mathbf{x}(t)+\mathbf{B}\omega_s(t)+\mathbf{E}\dot{f}(t) \qquad (9)$$

gde je $\mathbf{x}(t)=[e_d(t)\ \dot{e}_d(t)\ f(t)]^T$ vektor stanja, pri čemu matrice u modelu imaju formu:

$$\mathbf{A}=\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix},\ \mathbf{B}=\begin{bmatrix}0\\v_s\cos(\theta_e(t))\\0\end{bmatrix},\ \mathbf{E}=\begin{bmatrix}0\\0\\1\end{bmatrix} \qquad (10)$$

Na osnovu modela (9) linearni prošireni opserver stanja se projektuje na osnovu relacija:

$$\dot{\hat{\mathbf{x}}}(t)=\mathbf{A}\hat{\mathbf{x}}(t)+\mathbf{B}\omega_s(t)+\mathbf{L}(y-\hat{e}_d) \qquad (11)$$

gde je $\hat{\mathbf{x}}(t)=[\hat{e}_d(t)\ \hat{\dot{e}}_d(t)\ \hat{f}(t)]^T$ vektor estimiranih stanja, a matrica $\mathbf{L}=[\beta_1\ \beta_2\ \beta_3]^T$ sadrži pojačanja opservera. Koristeći estimacije stanja sistema i totalnog poremećaja, uz učešće merene veličine θ_e, može se formirati upravljački zakon:

$$\omega_s=\begin{cases}-\omega_{s\max}\ , \ \theta_e\le-\pi/2\\[2mm] \dfrac{-k_1\hat{e}_d-k_2\hat{\dot{e}}_d-\hat{f}}{v_s\cos(\theta_e(t))}\ ,\ |\theta_e|<\pi/2\\[2mm] \omega_{s\max}\ ,\ \theta_e\ge\pi/2\end{cases} \qquad (12)$$

gde je $\omega_{s\max}$ maksimalna ugaona brzina vozila. Uz pretpostavku da su estimacije zadovoljavajuće, tj. $\hat{e}_d(t)\approx e_d(t),\ \hat{\dot{e}}_d(t)\approx\dot{e}_d(t),\ \hat{f}(t)\approx f(t)$, primenom upravljačkog zakona (12) se kompenzuje uticaj totalnog poremećaja, a nelinearna dinamika greške (5) se svodi na zadatu formu (6), gde se dinamika približavanja vozila zadatoj putanji podešava parametrima k_1 i k_2.

Izbor parametara k_1 i k_2 se može izvršiti primenom metode podešavanja polova. Ako se za oba pola sistema sa zatvorenom spregom usvoji da su realni i isti, odnosno $p_{1,2}=-\omega_c$, $(\omega_c>0)$, pojačanja regulatora se mogu odrediti iz jednakosti karakterističnih polinoma:

$$(s+\omega_c)^2=s^2+k_2s+k_1\ , \qquad (13)$$

gde je sa ω_c definsan propusni opseg sistema sa regulatorom (12). Zadavanjem propusnog opsega opservera $\omega_o=k\omega_c$, gde je $k>1$, pojačanja opservera se mogu odrediti na sličan način, iz jednakosti karakterističnih polinoma:

$$(s+\omega_o)^3=s^3+\beta_1s^2+\beta_2s+\beta_3. \qquad (14)$$

Imajući u vidu mehanizam upravljanja kratnjem besposadnog vozila, treba napomenuti da su linerana brzina v_s i ugaona brzina rotacije vozila ω_s rezultat ugaonih brzina pogonskih točkova. Model kretanja guseničnog vozila sa dve gusenice i dva pogonska točka (za levu i desnu gusenicu) može se predstaviti u obliku:

$$\begin{aligned}v_s&=\frac{r}{2}(w_D+w_L)\\ \omega_s&=\frac{r}{m}(w_D-w_L)\end{aligned} \qquad (15)$$

gde su r prečnik pogonskog točka, m normalno rastojanje između gusenica, w_D i w_L ugaone brzine desnog i levog pogonskog točka, respektivno. Ukoliko uzmemo u obzir poremećaje usled proklizavanja gusenica, model (15) se može zapisati u obliku:

$$\begin{aligned}v_s+v_d&=\frac{r}{2}(a_Dw_D+a_Lw_L)\\ \omega_s+\omega_d&=\frac{r}{m}(a_Dw_D-a_Lw_L)\end{aligned} \qquad (16)$$

gde se koeficijentima a_D i a_L modeluje klizanje desne i leve gusenice, respektivno. Koeficijenti a_D i a_L su u opsegu $[0,1]$, pri čemu se za vrednosti manje od 1, generišu poremećaji v_d i ω_d. Odgovarajuće brzine pogonskih točkova w_D i w_L se direktno proračunavaju na osnovu zadatih vrednosti v_s i ω_s primenom izraza (15).

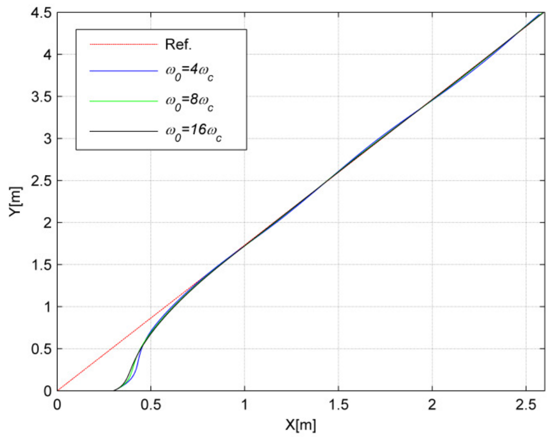
IV. SIMULACIONA ANALIZA

Simulaciona verfifikacija predloženog rešenja autonomnog kretanja je realizovana u programskom paketu MATLAB/Simulink na dinamičkom modelu vozila sa parametrima $m=1.4\,\mathrm{m}$ i $r=0.1\,\mathrm{m}$.

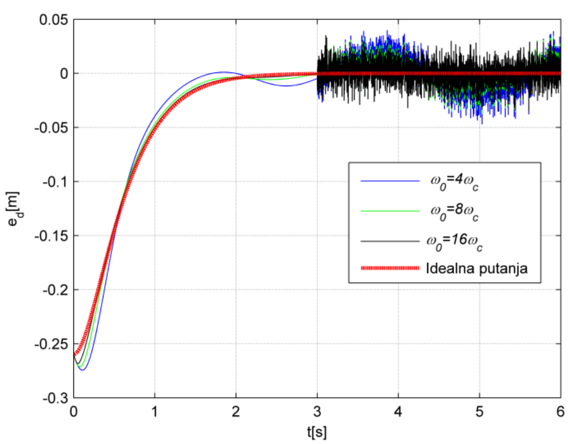
U okviru prvog scenarija razmatrano je kretanje vozila po zadatoj trajektoriji, definisanoj uglom $\theta_{ref}=\pi/3\,\mathrm{rad}$. Početni položaj vozila je definisan koordinatama $x(0)=0.3\,\mathrm{m}$, $y(0)=0\,\mathrm{m}$ i uglom $\theta_s(0)=\pi/4\,\mathrm{rad}$. Pretpostavljeno je postojanje proklizavanja obe gusenice, odnosno:

$$\begin{aligned}a_D(t)&=0.85+0.15\sin(2t+\pi/4)\\ a_L(t)&=0.85+0.15\cos(3t)\end{aligned}\ . \qquad (17)$$

Pored toga simulirano je postojanje šuma merenja signala $e_d(t)$, koji je uključen od 3 sekunde simulacije. Dobijeni rezultati praćenja zadate trajektorije, za tri ADRC regulatora (12) sa istim propusnim opsegom u zatvorenoj sprezi $\omega_c = 3$rad/s i različitim vrednostima propusnog opsega opservera ($\omega_o = 4\omega_c$, $\omega_o = 8\omega_c$ i $\omega_o = 16\omega_c$), su prikazani na Sl. 2. Kakakteristike praćenja u odnosu na idelanu putanju definisanu izrazom (6) su dati na Sl. 3, dok su vrednosti zahtevanih upravljačkih signala, odnosno brzina desnog i levog točka prikazane na Sl. 4. Kvalitet estimacija totalnog poremećaja za analizirane vrednosti propusnog opsega opservera je prikazan na Sl. 5.

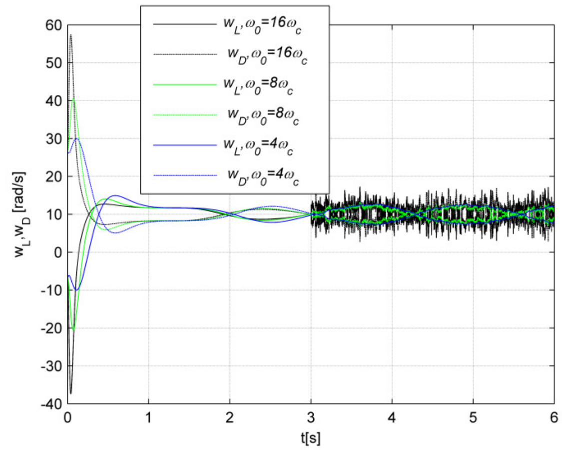


Sl. 2. Karakteristike praćenja zadate putanje za različite vrednosti propusnog opsega opservera

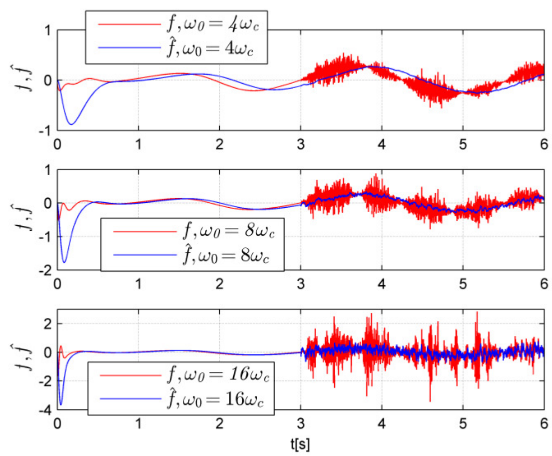


Sl. 3. Uporedne karakteristike praćenja u odnosu na idelanu putanju za različite propusne opsege opservera

Na osnovu rezultata sa Sl. 2 i Sl. 3 vidimo da vozilo, za sve tri vrednosti propusnog opsega opservera, uspešno prati zadatu putanju u uslovima postojanja pormećaja. Kao što je i očekivano, sistem sa najvećim propusnim opsegom opservera ostvaruje putanju najpribližniju idealnoj, što je posledica najmanje greške u estimaciji totalnog poremećaja (Sl. 5). Međutim, sa Sl. 4 se može uočiti da povećanje propusnog opsega opservera dovodi do porasta zahtevanih ugaonih brzina pogonskih točkova, kao i povećanja osetljivosti sistema na merni šum.



Sl. 4. Signali upravljanja pogonskim točkovima za različite vrednosti propusnog opsega opservera.

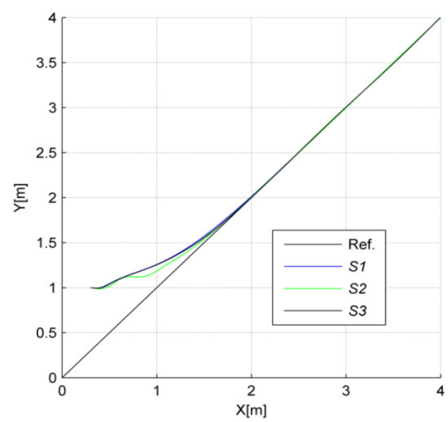


Sl. 5. Estimacije totalnog poremećaja za različite vrednosti propusnog opsega opservera.

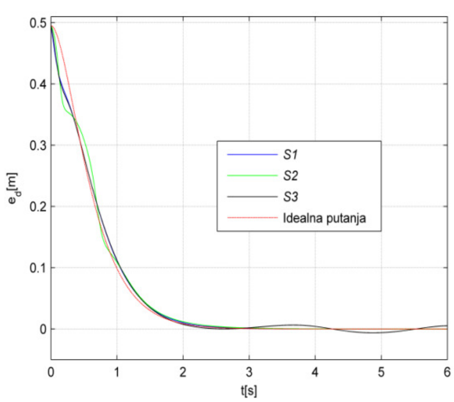
U drugom scenariju analizirano je praćenje putanje definisane uglom $\theta_{ref} = \pi/4$rad. Početni položaj vozila je definisan koordinatama $x(0) = 0.3$m, $y(0) = 1$m i $\theta_s(0) = 0\,rad$. Parametri regulatora su podešeni usvajanjem $\omega_c = 3\,rad/s$ i $\omega_o = 8\,\omega_c$, a simulirana su tri slučaja za različite dinamike proklizavanja gusenica. U prvom ($S1$) pretpostavljeno je da nema proklizavanja, odnosno, $a_D = 1, a_L = 1,$ dok je u druga dva slučaja proklizavanje modelovano sa:

$$S2: a_D = 0.85, a_L = 0.95;$$
$$S3: a_D(t) = 0.85 + 0.15\sin(2t + \pi/4), a_L(t) = 0.85 + 0.15\cos(3t);$$

Trajektorije vozila, greške praćenja normalne na trajektoriju $e_d(t)$ i estimacije totalnog poremećaja su prikazane na Sl. 6, Sl. 7 i Sl. 8, respektivno. Na osnovu dobijenih rezultata može se zaključiti da projektovani sistem ostvaruje zadovoljavajuće performanse praćenja u sva tri razmatrana slučaja. Kao što je i očekivano, najbolje performanse se dobijaju kada nema proklizavanja ($S1$) a vrednost totalnog poremećaja nula. U slučaju kada je proklizavanje gusenica konstantno ($S2$), putanja vozila odstupa od idealne u toku prilaska zadatoj trajektoriji, a nakon toga greška praćenja postaje nula, što je rezultat nulte greške estimacije totalnog poremećaja u stacionarnom stanju (Sl. 8). U trećem slučaju ($S3$) se uočava da, usled sinusoidalnog oblika proklizavanja, primenjeni prošireni opserver stanja ne može da estimira totalni poremećaj bez greške u stacionarnom stanju, (vidi Sl. 7), što dovodi do izvesnog odstupanja od zadate putanje.



Sl. 6. Karakteristike praćenja zadate putanje za različite dinamike proklizavanja gusenica.



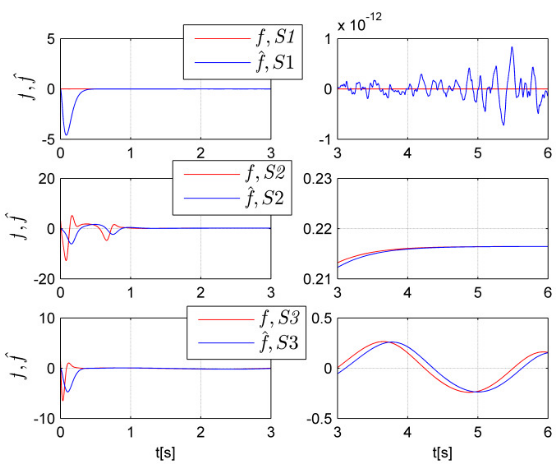
Sl. 7. Uporedne karakteristike praćenja u odnosu na idelanu putanju za različite dinamike proklizavanja gusenica.

V. Zaključak

Autonomno kretanje besposadnog vozila po zadatoj putanji, uz poremećaje linearne i ugaone brzine kretanja, realizovano je primenom kontrolera sa aktivnim potiskivanjem poremećaja. Problem je formulisan u vidu minimizacije najkraćeg (normalnog) rastojanja vozila od zadate putanje. Za estimaciju totalnog poremećaja projektovan je linearni prošireni opserver stanja, pri čemu je totalni poremećaj definisan tako da obuhvati i poremećaje koji ne deluju na istom ulazu kao i upravljački signal. Upravljački signal za aktivno potiskivanje totalnog poremećaja i približavanje vozila zadatoj putanji prema unapred definisanoj dinamici kretanja, je formiran na osnovu estimacija opservera i dostupnih merenja.

Koristeći model kretanja guseničnog vozila sa dve gusenice i dva pogonska točka, pokazano je da se poremećaji linearne i ugaone brzine vozila mogu efikasno estimirati primenom linearnog proširenog opservera stanja, pri čemu se povećanjem propusnog opsega opservera tačnost praćenja poboljšava, ali se povećavaju energetski zahtevi za aktuatore i osetljivost na merni šum. Primenjenim zakonom upravljanja uspešno se elimišu poremećaji tipa početnog stanja i konstantni poremećaji uzrokovani proklizavanjem, dok se uticaj dinamičnijih poremećaja može značajno umanjiti, uz odgovarajuće podešavanje parametara kontrolera.

Ideje za budući rad su, da se pre praktične implementacije na konkretnom besposadnom vozilu, analizira primena ADRC kontrolera za upravljanje intenzitetom brzine i ugaonom pozicijom vozila, sa ciljem autonomnog praćenja složenijih putanja.



Sl. 8. Estimacije totalnog poremećaja za različite dinamike proklizavanja gusenica (prelazni period - levo, stacionarno stanje - desno).

Zahvalnica

Rad je podržan od strane Ministarstva odbrane Republike Srbije u okviru Projekta VA/TT/1/21-23.

Literatura

[1] Y. Liu, R. Bucknall, "A survey of formation control and motion planning of multiple unmanned vehicles", *Robotica*, vol. *36*, no. 7, pp. 1019-1047, 2018.

[2] J. Shin, D. Kwak, T. Lee, "Robust path control for an autonomous ground vehicle in rough terrain", *Control Engineering Practice*, vol. *98*, 104384, 2020.

[3] J. Ni, J. Hu, C. Xiang, "A review for design and dynamics control of unmanned ground vehicle", *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. *235*, no. 4, pp. 1084-1100, 2021.

[4] Y. Huang, W. Xue, "Active disturbance rejection control: methodology and theoretical analysis". *ISA transactions*, vol. *53*, no. 4, pp. 963-976, 2014.

[5] J. Han, "From PID to active disturbance rejection control", *IEEE transactions on Industrial Electronics*, vol. 56, no. 3, pp. 900-906, 2009.

[6] C. H. E. N. Sen, X. U. E. Wenchao, L. I. N. Zhiyun, Y. Huang, "On active disturbance rejection control for path following of automated guided vehicle with uncertain velocities", In *2019 American Control Conference (ACC)*, pp. 2446-2451, IEEE, July, 2019.

Abstract

In this paper path following active disturbance rejection controller (ADRC), for unmanned ground vehicle (UGV), is proposed. Model nonlinarities, together with linear and angular velocity uncertainties, are treated as one total disturbance, enhancing both, matched and mismatched disturbances. Linear extended state observer (ESO) for total disturbance and nominal model states estimation is designed. Based on estimations and available measurements, the control signal for path following with predetermined transient behavior, is formulated. Simulation analysis, with tipical tracked unmanned ground vehicle motion configuration, is performed, and results illustrate efficiency of proposed control scheme in diferent path following scenarios.

Autonomous Active Disturbance Rejection Path Following Control of Unmanned Ground Vehicle

Momir Stanković i Stojadin Manojlović

Системи за подршку одлучивању базирани на вештачкој интелигенцији у третману умерене форме билијарног панкреатитиса

Ања Буљевић, Александар Глуховић, Мирна Н. Капетина, Александар Кнежевић, Зоран Д. Јеличић

***Апстракт*—У оквиру овог рада представљено је једно решење система за подршку одлучивању у лечењу умерене форме билијарног панкреатитиса. На основу параметара и показатеља из клиничке праксе, развијен је систем за усмеравање лекара приликом избора методе и поступка лечења овог запаљења панкреаса. Због релативно малог скупа података, услед специфичних медицинских процедура, избор релевантних обележја потврђен је кроз два формализма: корелационом анализом и стаблом одлучивања. На основу одабраних обележја, до финалног решења се долази уз ослонац на теорију потпорних вектора. Предложено решење нашло је своју примену у клиничкој пракси.**

***Кључне речи*: умерена форма билијарног панкреатитиса, корелациона анализа, стабло одлучивања, СВМ.**

I. Увод

У наставку текста представићемо основне показатеље и параметре који карактеришу умерену форму билијарног панкреатитиса, али и конвенционалне и не тако конвенционалне начине лечења истог. Наиме, полазећи од претпоставке да се ради о мултидисциплинарном проблему са чистом медицинском применом разумевање медицинског дела је од највећег значаја не само за крајње кориснике, већ као и тумачење изабраних обележја и разумевања свих сложених феномена који се могу добити у математичком опису проблема.

Акутни панкреатитис представља ензимско инфламаторно обољење панкреаса, које може захватити како сам орган, тако и околна ткива. Инциденција обољења је око 17/100000 становника. Најчешћи етиолошки чиниоци који се везују за ово стање су билијарна калкулоза (45%) и конзумација алкохолних пића (35%). Умерене форме акутног панкреатитиса, које су и предмет овог истраживања, јављају се у 80% случајева и имају благ клинички ток са стопом морталитета од 1%. Акутни панкреатитис билијарне етиологије узрокован је калкулозом жучне кесе и/или жучних путева. Третман калкулозе билијарног стабла код умерене форме панкреатитиса, по актуелној препоруци Америчког удружења гастроентеролога и ендоскопских хирурга (American Assosiation of Gastroenterology and Endoscopic Surgeons - SAGES), своди се на уклањање жучне кесе лапароскопском холецистектомијом (ЛХ) са интраоперативном холангиографијом (ИОХ), у циљу превенције појаве новог атака болести. Уколико постоји сумња на присуство калкулуса у жучним каналима са повишеним вредностима билирубина, ради се ендоскопска ретроградна холангиопанкреатографија (ЕРЦП) са ендоскопском папилотомијом (ЕПТ), у циљу уклањања калкулуса и детритуса из жучних водова, и обезбеђивања нормалног протока жучи у дванаестопалачно црево. Ова процедура се углавном изводи пре ЛХ, мада се може радити и током ЛХ или након ње, [1].

На студији случајева Војводине, ЛХ је доступна у свим болницама, као и у клиничком центру. Међутим, ЕРЦП процедура могућа је само у Клиничком центру Војводине и то од стране једног лекара. По нашим најбољим сазнањима, укупан број лекара у Србији који изводе ову операцију је 3 или 4. Као кључно питање и задатак намеће се одабир скупа објективних параметара који ће лекара определити да само изведе широко доступну ЛХ или да ипак пацијента пошаље и на додатну процедуру ЕРЦП. Важно је напоменути да су анализирани само објективне параметре јер субјективни параметри (нпр. ултразвучни преглед) не могу да гарантују доследност у резултатима. Други начини прегледа, попут компјутеризоване томографије (ЦТ) и магнентна резнонце, нису широко доступни. При томе, за око 80% пацијената довољна је ЛХ, док свега 20% пацијената захтева и ЕРЦП.

Кључно за анализе у систему за подршку одлу-

А. Буљевић (anjabuljevic@uns.ac.rs)*, А. Глуховић (aleksandar.gluhovic@gmail.com)**, М. Н. Капетина (mirna.kapetina@uns.ac.rs)*, А. Кнежевић (aknezevic021@gmail.com)***, З. Д. Јеличић (jelicic@uns.ac.rs)*
* Универзитет у Новом Саду, Факултет техничких наука, Департман за рачунарство и аутоматику, Трг Доситеја Обрадовића 6, 21000, Нови Сад, Р. Србија
** New Hospital, Алберта Ајнштајна, Нови Сад, Р. Србија
*** Клинички центар Војводине, Хајдук Вељкова 1-9, Нови Сад, Р. Србија

чивању је било да минимизујемо лажно негативне закључке у потреби извођења ЕРЦП што је и заиста у складу са добром клиничком праксом. Због малог скупа података, избор обележја је урађен кроз корелациону анализу и стабло одлучивања што је оправдано са математичког аспекта [2], а такав начин рада се користи и у сличним медицинским истраживањима [3]. Према нашим најбољим сазнањима, овакав приступ подршке одлучивању у лечењу умерене форме акутног панкреатитиса је оригиналан и први пут се презентује у овом раду.

Рад је организован на следећи начин. У поглављу II приказан је одабир дискриминантних обележја. У поглављу III приказано је балансирање података. Формирање система за подршку одлучивању дато је у поглављу IV, док је закључак дат у поглављу V.

Ова студија је одобрена од стране Етичког комитета Медицинског факултета у Новом Саду и Етичке комисије Клиничког центра Војводине.

II. КОРЕЛАЦИОНА АНАЛИЗА И СТАБЛО ОДЛУЧИВАЊА

Први корак за решавање овог реалног проблема била је идентификација дискриминантних обележја (медицинских параметара) који су од круцијалне важности за указивање присуства калкулуса у жучним водовима. За потребе овог пројекта коришћени су медицински подаци пацијаната оболелих од акутног панкреатитиса који су прикупљени у Клиничком центру Војводине. Приликом самог пријема пацијената прикупљено је преко 80 параметара. Неки од њих су били дескриптивни подаци о субјективном стању пацијената, затим о историји болести, као и објективни параметри попут налази крви и субјективни као што је ултразвучни преглед. Сви параметри су по правилу распоређени у временским серијама од 12 часова у првих 48 сати по пријему у болницу.

У добијеној бази података налазе се подаци за 100 пацијената лечених од умерене форме билијарног панкреатитиса. У литератури [1] проналазимо да се потенцијални кандидати за ЕРЦП могу узети у разматрање уколико је њихов Глазгов скор мањи од 3. Због ове чињенице, база се смањује на 96 пацијента. Од тих 96 пацијената, код њих 19 је рађен ЕРЦП, док је код 77 пацијената била довољна ЛХ.

Кроз нумеричке поступке издвајања обележја и пратећи начин размишљања лекара издвојили смо 5 објективних параметара: укупни билирубин, директни билирубин, алкална фосфатаза, ЦРП и гамаГТ. Субјективни параметри су показали извесну недоследност и нису могли бити узети у разматрање. Испитивање дискриминантности обележја рађено је на два начина:

1) коришћењем Пирсоновог коефицијента корелације и
2) коришћењем стабла одлучивања.

А. Одабир обележја коришћењем Пирсоновог коефицијента корелације

Поступак идентификације дискриминантних обележја започели смо корелационом анализом. Уколико желимо да испитамо да ли постоји зависност између два (или више) обележја, тада говоримо о утврђивању постојања корелације између тих обележја [4]. Коефицијент корелације је показатељ степена статистичке повезаности обележја и представља меру њихове линеарне зависности. Једна од најчешће коришћених мера повезаности два обележја јесте Пирсонов коефицијент корелације [4]. У табели I су приказани Пирсонови коефицијенти корелације за она обележја која имају највећи тражени коефицијент и они уједно представљају најзначајнија обележја за наш проблем.

Табела I
Пирсонов коефицијент корелација за дискриминантна обиљежја

Назив обележја	$\|\rho\|$
укупни билирубин	0.4051
директни билирубин	0.3837
алкална фосфатаза	0.3703
ЦРП	0.274
гама ГТ	0.2164

Из ове табеле можемо приметити да су чак и најзначајнија обележја по Пирсоновом коефицијенту корелације уствари у слабој корелацији са излазном променљивом. Како бисмо потврдили да су одабрана обележја заиста дискриминантна за наш проблем, урађена је и валидација добијених обележија коришћењем стабла одлучивања.

Б. Одабир обележја коришћењем стабла одлучивања

Други начин одређивања дискриминантних обележја који је имплементиран у овом раду јесте стабло одлучивања. Стабло одлучивања представља графички модел за визуализацију процеса одлучивања када се решавање проблема одлучивања своди на доношење више сукцесивних одлука, [5]. На почетку се бира параметар чија вредност најбоље дели расположиве узорке. Као што је познато, стабло одлучивања се осим за одређивање дискриминантних обележја, може користити и за класификацију података. Због малог броја података, као и слабе корелације улазних параметара са излазом, у овом случају стабло одлучивања је коришћено искључиво за селекцију обележја, док ће се метода вектора носача користити као подршка одлучивању у третману умерене форме билијарног панкреатитиса. Стабло одлучивања, поред селекције, омогућава нам и увид у значајност изабраних обележја, као и сам „пут” селекције. Предност стабла одлучивања у односу на Пирсонов коефицијент корелације јесте аутоматска селекција обележја [6].

За потребе овог рада коришћен је ID3 алгоритам. Квантитативна мера коју ID3 алгоритам користи како би одредио најбоља обележја јесте информациона добит [7]. Симулације су вршене са различитим конфигурацијама стабла. Дубина стабла је узимала вредности на интервалу од 3 до 10, док се број листова се налазио у опсегу од 2 до 10. Посматрајући добијена стабла одлучивања примећено је да се као заједнички садржаоци свих стабала издвајају следећа обележја: укупни билирубин, ЦРП, директни билирубин, алкална фосфатаза и гама ГТ. Осим тога, наведена обележја се у већини генерисаних стабала налазе ближе корену стабла, односно на мањим дубинама стабла. Упоређујући резултате које смо добили коришћењем стабла одлучивања и Пирсоновог коефицијента корелације, потврдили смо потпуну подударност између дискриминантних обележја добијених коришћењем ова два поступка.

III. БАЛАНСИРАЊЕ ПОДАТАКА

Следећи проблем који се намеће је небалансираност података. Небалансираност података подразумева ситуацију у којој се број узорака значајно разликује по класама у поступку класификације. Класификатори машинског учења тешко се носе са небалансираним скупом података за обуку, јер су осетљиви на пропорционалност различитих класа, па се као последица јавља тенденција алгоритама да фаворизују класу са највећим уделом испитаника што најчешће резултује „обмањујућом” тачношћу. То је посебно проблематично када је од интереса тачна класификација „ретке” (мањинске) класе, али налазимо висок проценат тачности који је заправо последица исправне класификације већинске класе. Ову чињеницу овде експлицитно наводимо, не само да нагласимо потребу за балансирањем података, већ и да уведемо посебну метрику за оцену квалитета предикције. С обзиром на то да алгоритми машинског учења имају за циљ да смање укупну стопу грешке, неће обраћати посебну пажњу на мањинску класу, и вероватно неће успети да направе тачно предвиђање за ову класу, јер о њој не садржи довољно података.

У бази података разматраног проблема, од укупно 96 пацијената, код само 19 испитаника је рађен ЕРЦП, док код 77 испитаника није рађен ЕРЦП. Као што је наведено у уводном поглављу, приоритет је да се минимизују лажно негативни закључци у потреби извођења ЕРЦП који у овом случају представља мањинску класу. Дакле, неопходно је било урадити балансирање података.

У литератури [8] се може пронаћи неколицина потенцијалних метода за решавање овог проблема, а као најбољи метод се наводи додавање нових података у класе са процентуално мањим бројем узорака. Међутим, у пракси је то веома тешко постићи, па се прибегава неким другим методама. Један од најчешће коришћених је да се из већинске класе избаце узорци како би се број узорака већинске класе изједначио са бројем узорака мањинске класе. При томе, мора се водити рачуна да укупан број узорака мора бити бар 10 пута већи од броја изабраних обележја [9].

Ова препорука за балансирање података је искориштена у овом раду на следећи начин. На случајан начин се од 77 испитаника код којих није рађен ЕРЦП изабере 40 и тих 40 испитаника улазе у процес обуке СВМ алгоритмом, заједно са 19 испитаника код којих је рађен ЕРЦП. Примећујемо да подаци и даље нису најбоље балансирани, али због ограничења да укупан број узорака мора бити барем 10 пута већи од укупног броја обележја, ово је најбоље што смо могли да добијемо из кориштене базе података.

IV. СВМ МАТЕМАТИЧКИ МОДЕЛ

Након што смо идентификовали параметре од интереса и избалансирали податке, неопходно је било да се на основу издвојених параметара формира систем за подршку одлучивању у третману умерене форме акутног билијарног панкреатитиса. Због заиста малог скупа података, нарочито малог за пацијенте којима је рађен ЕРЦП, определили смо се за математички модел уз ослонац на теорију потпорних вектора (СВМ) [10]. Метода класификације базирана на векторима носачима представља један од модела машинског учења који се веома често користи како за класификациону, тако и за регресиону анализу. СВМ алгоритам је довољно познат алгоритам, па неће бити детаљно извођен у овом раду, а његово детаљно математичко извођење можете пронаћи у литератури [11], [12].

Посебна пажња је била посвећена оптимизацији параметара СВМ алгоритма. Будући да перформансе генерализације СВМ алгоритма у великој мери зависе од параметара C (хиперпараметар који прави компомис између сложености модела и степена до кога се толеришу одступања модела), ε (хиперпараметар који контролише ширину неосетљиве зоне, а његова вредност утиче на број вектора носача) и γ (параметар који одређује облик изабране кернел функције), неопходно је извршити њихову оптимизацију. Према [13], [14], између наведених параметара постоји јака веза, тако да је препорука да се они оптимизују истовремено, а не одвојено. Оптимизација параметара је извршена уз помоћ алгоритма роја честица [15] и као критеријум оптималности је кориштена средња квадратна грешка [16]. Као кернел функције су прослеђиване: радијална, гаусова, линеарна и полиномна кернел функција. Након оптимизација кернел функција и свих њених пратећих параметара, добијено је да се најбољи резултати добију за СВМ алгоритам који има радијалну кернел функцију и добијени су следећи параметри: $C = 10^2$, $\varepsilon = 10^{-3}$ и $\gamma = 10^{-1}$.

Иако је урађено балансирање података, класе и даље нису у потпуности балансиране. Осим тога, СВМ модел је јако осетљив на улазне податке и неретко од скупа улазних података, зависи и тачност класификације СВМ. Како би се добили што објективинији резултати, спроведен је експеримент описан у наставку. Узорци који се прослеђују СВМ као улазни подаци су прослеђивани на следећи начин:

- од 19 испитаника код којих је рађен ЕРЦП, на случајан начин се бира 80% испитаника (15 испитаника) који се прослеђују СВМ алгоритму за обуку и валидацију, док је на преостала 4 узорка вршено тестирање добијеног модела
- од 77 испитаника код којих није рађен ЕРЦП на случајан начин се изабере 40 испитаника. Од тих 40 испитаника, на случајан начин се бира 80% испитаника (32 испитаника) који се прослеђују СВМ алгортиму за обуку и валидацију, док је на преосталих 8 узорака вршено тестирање добијеног модела.

Као оцена успешности одабраног класификационог модела, уобичајено се користе стандардизоване мере и оцене којима се квантификује рад пројектованог система за класификацију и предикцију. Под оцењивањем, односно процењивањем рада система за предикцију углавном се мисли на одређивање вредности неких од стандардних мера којима се квантификује његов учинак односно перформансе. Мера квалитета представља потенцијал модела да коректно предвиди класу новог податка. Матрица конфузије (engl. *confusion matrix*) представља детаљан и прегледан приказ бројева исправно и погрешно класификваних узорака на основу којих се могу вршити оцене добијеног модела класификације. Општи облик матрице конфузије за бинарну класификацију је приказан табелом II, где је

- TP (true positive; стварно позитивни) - број узорака који припадају позитивној класи, а додељена им је позитивна класа,
- TN (true negative; стварно негативни) - број узорака који припадају негативној класи, а додељена им је негативна класа,
- FP (false positive; лажно позитивни) - број узорака који припадају негативној класи, а додељена им је позитивна класа,
- FN (false negative; лажно негативни) - број узорака који припадају позитивној класи, а додељена им је негативна класа

Табела II
Општи облик матрице конфузије.

		Предвиђена класа	
		Класа = 0	Класа = 1
Праве класе	Класа = 0	TN	FP
	Класа = 1	FN	TP

Табела III представља један пример матрице конфузије СВМ класификатора за случај када треба да се спроведе ЕРЦП. Примећујемо да је у овом случају од 40 испитаника којима је довољна само ЛХ процедура, наш класификациони модел то погодио за 39 испитаника, док је за само једног испитаника рекао да му је неопхдан и ЕРЦП. Што се тиче 19 испитаника којима је неопходан и ЕРЦП, наш класификациони модел је одговарајућу класу погодио за 17 испитаника, док је за двојицу испитаника погрешно доделио класу.

Табела III
Матрица конфузије за случајеве када је неопходан ЕРЦП.

		Предикција	
		ЛХ	ЛХ и ЕРЦП
Стварни резултати	ЛХ	39	1
	ЛХ и ЕРЦП	2	17

По завршетку креирања класификационог модела, корисно је тестирати његове перформансе на скупу података који му је непознат, при чему је неопходно да тај скуп података садржи информације о класама. Овакав вид тестирања представља непристрасну оцену генерализације. Најчешће се користи К-слојна унакрсна валидација (енг. K-fold cross-validation). К-слојна унакрсна валидација (кросвалидација) је техника евалуације класификационих модела која се изводи тако што се оригинални скуп података дели на k једнаких подскупова. Један подскуп се користи за тестирање, док се сви остали подскупови користе за тренирање. Овај поступак се понавља у k итерација тако да се сваки подскуп користи тачно једном за тестирање. По завршетку свих итерација, издваја се онај модел који је имао најмању грешку класификације, [17].

Већ је напоменуто да је излаз јако осетљив на добијени скуп улазних података, па да би се добили што објективнији резултати поступак описан горе је поновљен у 100 итерација са одабраним СВМ моделом. На овај начин су праћене две тачности модела: тачност над обучавајућим скупом података у процесу кросвалидације и тачност над тестним скупом. Тачност, као једна од најчешће коришћених мера за приказивање успешности класификације, представља однос укупног броја коректних предвиђања и укупног броја предвиђања. Математички, тачност записујемо

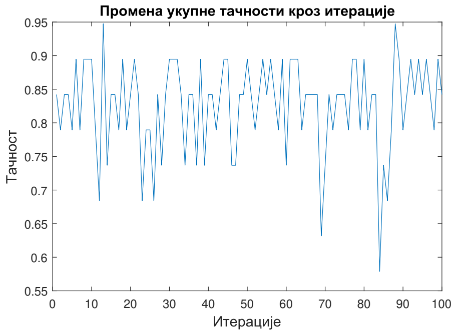
$$\text{тачност} = \frac{TP + TN}{TP + TN + FP + FN}.$$

Промена тачности над обучавајућим скупом у процесу кросвалидације за одабрани модел за 100 итерација у зависности од одабраних улазних података дат је на слици 1. Просечна тачност овог модела износи 90.51%.

Промена тачности над тестним скупом за одабрани модел за 100 итерација у зависности од одабраних улазних података дата је на слици 2. Просечна тачност овог модела износи 82.68%.

Поред тачности, морамо обратити пажњу на још неколико показатеља успешности класификације. Прецизност је мера слична тачности, али се односи искључиво на једну посматрану класу. Она представља

Слика 1. Промена укупне тачности у процесу кросвалидације кроз итерације



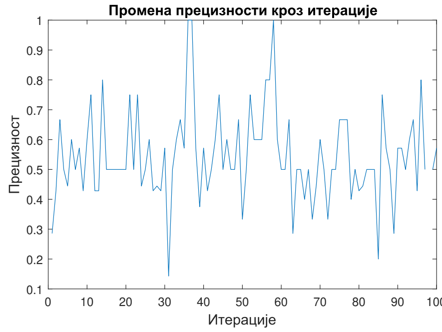
Слика 2. Промена укупне тачности кроз итерације

однос тачних позитивних предвиђања и укупног броја случајева у којима је класификатор предвидио посматрану класу. Осетљивост приказује однос коректно предвиђених вектора атрибута неке класе и укупног броја правих понављања те класе у скупу података. Специфичност је способност теста да коректно идентификује одсуство неког атрибута, а може да се интерпретира и као процена условне вероватноће да атрибут није идентификован, уз услов да га на посматраној позицији заиста нема. Наведене мере математички можемо записати на следећи начин

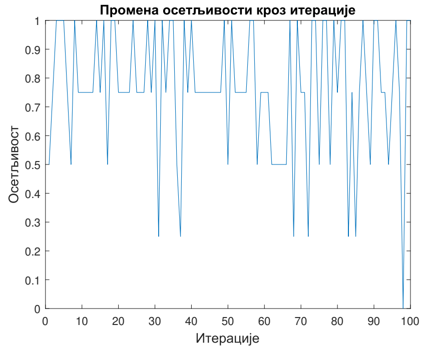
$$\text{прецизност} = \frac{TP}{TP+FP},$$
$$\text{осетљивост} = \frac{TP}{TP+FN},$$
$$\text{специфичност} = \frac{TN}{FP+TN}.$$

Графици наведених мера прецизности, осетљивости и специфичности су приказани на сликама 3, 4 и 5 респективно. Просечна вредност прецизности износи 73.23%, просечна вриједнoст осетљивости износи 76.25%, док просечна вредност специфичности износи 64.88%.

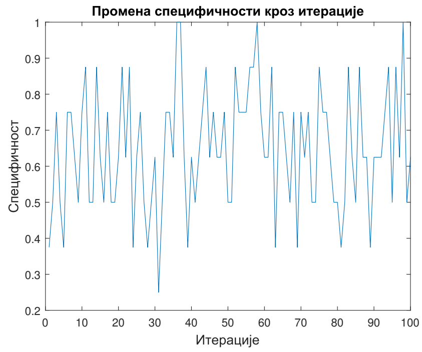
Као што је већ речено, нама је од посебне важности тачност погађања за случајеве када је непоходна ЕРЦП процедура, односно битно нам је да одговор класификационог алгоритма за случај када треба да се ради ЕРЦП процедура буде тачан. Због тога је посебно издвојена и тачност модела само за ову класу и приказана је на слици 6. Просечна вредност тачности за овај случај износи 76.25%.



Слика 3. Промена прецизности кроз итерације



Слика 4. Промена осетљивости кроз итерације



Слика 5. Промена специфичности кроз итерације

Слика 6. Промена тачности за случај када је неопходан ЕРЦП кроз итерације

Анализирајући вредности за приказане мере, можемо да закључимо да одабрани класификациони СВМ модел даје задовољавајуће резултате. Примећујемо да је укупна тачност модела већа у односу на тачност модела када испитујемо случај када је неопходна и ЕРЦП процедура, али то смо и очекивали пошто та класа представља мањинску класу за овај проблем.

V. ЗАКЉУЧАК

У склопу овог рада пројектован је систем за подршку одлучивању у третману умерене форме билијарног панкреатитиса. Пројектовању система за подршку одлучивању претходила је предобрада података. Приликом самог пријема пацијената прикупљено је преко 80 медицинских параметара, па је било неопходно одабрати најзначајније параметре који би нам указали на присисутво калкулозе у жучним водовима. Дискриминантна обележја су одређена коришћењем Пирсоновог коефицијента корелације и као дискриминантна обележја су се издвојили: укупни билирубин, ЦРП, директни билирубин, алкална фосфатаза и гама ГТ. Ваљаност издвојених обележја потврђена је применом стабла одлучивања.

Главни циљ нашег рада је био да минимизујемо лажно негативне закључке у потреби извођења ЕРЦП. За формирање система за подршку одлучивању коришћен је СВМ алгоритам, коме је претходила процедура балансирање података по класама. Оптимизација параматера модела над тестним скупом података је била 82.68%. Пошто нам је од посебне важности била тачна предикција случајева када је неопоходан ЕРЦП, издвојена је и просечна тачност модела за овај случај и она износи 76.25%.

Битно је напоменути да лекари у својој клиничкој пракси заиста користе издовјена обележја како би идентификовали присуство калкулузе, а последично следи одлука да ли треба да се ради додатна процедура како би се решио проблем оваквог оболења панкреаса. Оно што се математички не може описати, бар не на малом скупу података, јесте субјективни осећај, знање и искуство искусног лекара из клиничке праксе који приликом ЛХ може да примети да ли у жучним водовима постоји калкулуза и на основу тога да донесе процену да ли је потребан ЕРЦП.

ЗАХВАЛНИЦА

Овај рад је подржан од стране Министарства просвете, науке и технолошког развоја кроз пројекат број 451-03-68/2020-14/200156: „Иновативна научна и уметничка испитивања из домена делатности ФТН-а"

ЛИТЕРАТУРА

[1] А. Глуховић, "Алгоритам примене лапароскопске холецистектомије и ендоскопске ретроградне холангиопанкреатографије са папилотомијом у третману умерене форме билијарног панкреатитиса." Докторска дисертација. Нови Сад, 2016.

[2] H. Zhou, J. Zhang, Y. Zhou, X. Guo, and Y. Ma, "A feature selection algorithm of decision tree based on feature weight," *Expert Systems with Applications*, vol. 164, p. 113842, 2021.

[3] Q. Qiu, Y.-j. Nian, Y. Guo, L. Tang, N. Lu, L.-z. Wen, B. Wang, D.-f. Chen, and K.-j. Liu, "Development and validation of three machine-learning models for predicting multiple organ failure in moderately severe and severe acute pancreatitis," *BMC gastroenterology*, vol. 19, no. 1, pp. 1–9, 2019.

[4] M. A. Hall, "Correlation-based feature selection for discrete and numeric class machine learning," in *Proceedings of the Seventeenth International Conference on Machine Learning*, ICML '00, (San Francisco, CA, USA), p. 359–366, Morgan Kaufmann Publishers Inc., 2000.

[5] A. Geron, "Hands-on machine learning with scikit-learn, keras tensorflow." O'Reilly, 2019.

[6] K. G. (auth.), *Meta-Learning in Decision Tree Induction*. Studies in Computational Intelligence 498, Springer International Publishing, 1 ed., 2014.

[7] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees.* Monterey, CA: Wadsworth and Brooks, 1984.

[8] G. E. Batista, R. C. Prati, and M. C. Monard, "A study of the behavior of several methods for balancing machine learning training data," *ACM SIGKDD explorations newsletter*, vol. 6, no. 1, pp. 20–29, 2004.

[9] V. Sugumaran and K. Ramachandran, "Effect of number of features on classification of roller bearing faults using svm and psvm," *Expert Systems with Applications*, vol. 38, no. 4, pp. 4088–4096, 2011.

[10] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[11] B. Schölkopf, A. J. Smola, F. Bach, *et al.*, *Learning with kernels: support vector machines, regularization, optimization, and beyond.* MIT press, 2002.

[12] F. Dedov, *The Python Bible Volume 4: Machine Learning (Neural Networks, Tensorflow, Sklearn, SVM).* The Python Bible Series, Independently Published, 2019.

[13] V. Cherkassky and Y. Ma, "Practical selection of svm parameters and noise estimation for svm regression," *Neural networks*, vol. 17, no. 1, pp. 113–126, 2004.

[14] Z. Shaowu, W. Lianghong, Y. Xiaofang, and T. Wen, "Parameters selection of svm for function approximation based on differential evolution," in *International Conference on Intelligent Systems and Knowledge Engineering 2007*, pp. 529–535, Atlantis Press, 2007.

[15] Ž. Kanović, M. R. Rapaić, and Z. D. Jeličić, "Generalized particle swarm optimization algorithm-theoretical and empirical analysis with application in fault detection," *Applied Mathematics and Computation*, vol. 217, no. 24, pp. 10175–10186, 2011.

[16] Y. Ren and G. Bai, "Determination of optimal svm parameters by using ga/pso," *Journal of computers*, vol. 5, no. 8, pp. 1160–1168, 2010.

[17] C.-W. Hsu, C.-C. Chang, C.-J. Lin, *et al.*, "A practical guide to support vector classification," 2003.

Abstract

This paper presents one solution for the decision support system in the treatment of moderate form of biliary pancreatitis. Based on parameters and indications from clinical practices, the system for directing doctors in selecting methods and procedures for treating this pancreas ailment has been developed. Due to specific medical procedures, the data set is relatively small, so the choice of relevant features was confirmed through two formalisms: the correlation analysis and the decision tree. Based on the selected features, the final solution is reached by the theory of supporting vectors. The proposed solution has found its application in clinical practice.

Decision support system based on artificial intelligence in the treatment of moderate form of biliary pancreatitis

Anja Buljević, Aleksandar Gluhović, Mirna N. Kapetina, Aleksandar Knežević, Zoran D. Jeličić

Predikcija ishoda protetičke rehabilitacije nakon amputacije donjih ekstremiteta uz oslonac na algoritme veštačke inteligencije

Jovana Arsenović, Aleksandar Knežević, Mirna N. Kapetina i Zoran D. Jeličić

***Apstrakt*—Protetička rehabilitacija trenutno predstavlja najbolji tretman za pacijente sa amputacijom donjih ekstremiteta. Međutim, fabrikacija proteze i prateća protetička rehabilitacija predstavljaju veoma dug i skup proces koji nekada ne dovodi do poboljšanja mobilnosti i kvaliteta života pacijenata. Zbog toga je neophodno predvideti ishod rehabilitacionog tretmana. Glavni zadatak ovog rada bio je da se napravi alat, uz oslonac na algoritme veštačke inteligencije, koji bi se mogao primeniti u ranim fazama, kako bi se napravila što bolja predikcija ishoda rehabilitacije pacijenata sa amputacijom donjih ekstemiteta, odnosno kako bi se predvideo K-nivo (engl. *Medicare Functional Classification Level, K-level*), ishod testa dvominutnog hoda (engl. *two minute walk test*) i testa ustani i kreni (engl. *timed up and go test*). Evaluacija modela vršena je nad realnim podacima pacijenata Klinike za rehabilitaciju, Kliničkog centra Vojvodine. Dobijeni rezultati pokazuju značajno poboljšanje, u pogledu performansi klasifikatora, u odnosu na prethodne metode i potvđuju izbor nekih od najznačajnijih parametara prilikom identifikacije pacijenata.**

***Ključne reči*—Predikcija, veštačka inteligencija, metoda vektora nosača - SVM, stabla odluke, amputacija, rehabilitacija.**

I. Uvod

Amputacija donjih ekstremiteta predstavlja hirurški postupak koji se primenjuje radi odstranjivanja ishemičnog, inficiranog, nekrotičnog tkiva ili lokalnog tumora, kada nije moguća resekcija [1]. U svetu se godišnje izvrši preko milion amputacija noge, procena je Svetske zdravstvene organizacije i Međunarodne dijabetološke federacije, dok se prema podacima Kliničkog centra Vojvodine, u toj ustanovi godišnje izvrši više od 100 amputacija donjih ekstremiteta iznad nivoa skočnog zgloba [1]. Kretanje predstavlja osnovnu potrebu čoveka, a hod je primarni način kretanja ljudi, tako da je glavni cilj rehabilitacionog procesa ponovno uspostavljanje ove funkcije. Poboljšanje sveobuhvatnog stanja i kvaliteta života osoba sa amputacijom donjih ekstremiteta omogućava protetička rehabilitacija.

Proteze mogu omogućiti funkionalni hod i nadomestiti fizički nedostatak i trenutno predstavljaju najbolje rešenje za osobe sa amputacijom donjih ekstremiteta [2]. Nažalost, nisu sve osobe sa amputacijom dobri kandidati za protetičku rehabilitaciju, a ovaj skup i dug proces nekad ne dovodi do poboljšanja mobilnosti i kvaliteta života u meri u kojoj se očekivalo. Stoga postoji potreba da se predvidi ishod potencionalnog rehabilitacionog tretmana.

Cilj ovog istraživanja bio je da se napravi alat, uz oslonac na algoritme veštačke inteligencije, koji bi se mogao što ranije primeniti, kako bi se napravila što bolja predikcija ishoda rehabilitacije pacijenata sa amputacijom, odnosno kako bi se predvideo K-nivo i ishodi testa dvominutnog hoda i testa ustani i kreni. Da bi predikcija bila što uspešnija potrebno je identifikovati parametre za predikciju, odnosno one faktore koji utiču na osposobljenost za hod uz pomoć proteze. Prilikom identifikovanja pacijenata, odnosno tokom donošenja odluke da li je pacijent dobar kandidat za propisivanje proteze i započinjanje protetičke rehabilitacije, lekari se fokusiraju na određenje parametre (engl. *feature*, u daljem tekstu obeležje, atribut, faktor, varijabla ili parametar). Oni moraju biti strogo definisani, što jednostavniji za procenu, da ih ne bude previse, a ni premalo, tako da doprinose najboljoj tačnosti klasifikatora i ne predstaljaju problem lekarima za određivanje. Svi autori koji su pisali na ovu temu, složili su se da postoji više faktora koji utiču na uspešnost predikcije hoda uz pomoć proteze, ali koji su dominantni i dalje je predmet istraživanja.

U radu [3] za konstruisanje predikcionog modela baziranog na vektorima nosačima (engl. *support vector machine*, u daljem tekstu SVM) korišćeno je 11 varijabli: starost, pol, uzrok i nivo amputacije, period od amputacije do protetičke rehabilitacije, funkcionalni komorbidetni indeks, pisustvo šećerne bolesti, prisustvo partnera, ograničena ekstenzija kuka ili kolena rezidualnog ekstremiteta i mobilnost pri prijemu. Za izbor najboljih obeležja u radu [3] korišćen je genetski algoritam, koji je izdvojio starost, funkcionalni komorbidetni indeks, nivo amputacije i mobilnost pri prijemu kao dominantne faktore. Kriterijum optimalnosti bila je tačnost klasifikacije nad test skupom. Preciznost modela bila je u intervalu od 72,5% do 82,5%. Kao dominantna obeležja za klasifikaciju u radu [4] izabran je test šestominutnog hoda (engl. *6-minute walk test*) i test stojanja na jednoj nozi (engl. *one-leg standing test*). Na osnovu ovih parametara K-nivo su predviđali sa osetljivošću blizu 90%. Ono što su u ovom radu istakli jeste da je klasifikacija rađena samo za pacijente sa transtibijalnom amputacijom. U radu [5] kao nezavisni

Jovana Arsenović (email: arsenovic.jovana@uns.ac.rs), Mirna N. Kapetina (email: mirna.kapetina@uns.ac.rs), Zoran D. Jeličić (email: jelicic@uns.ac.rs) - Univerzitet u Novom Sadu, Fakultet tehničkih nauka, Katedra za automatsko upravljanje, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija.

Aleksandar Knežević (email: aleksandar.knezevic@mf.uns.ac.rs) – Univerzitet u Novom Sadu, Medicinski fakultet, Katedra za fizikalnu medicinu i rehabilitaciju, Hajduk Veljkova 3, 21137 Novi Sad, Srbija.

prediktivni faktori predloženi su pol pacijenta, starost, dužina trajanja rehabilitacije i dužina čekanja na početak primarne protetičke rehabilitacije, dok su u radu [6] izdvojena sledeća obeležja: pređenja razdaljina i trajanje hoda, nivo amputacije, pol i vremenski interval od amputacije do protetičke rehabilitacije. Trenutno, ne postoje tačno definisane kliničke preporuke za određivanje kandidata za protetičku rehabilitaciju, kao ni jasnih faktora koji bi potencijalno mogli predvideti ishod rehabilitacije. U radovima [7, 8] izvršen je sistematičan pregled literature napisan na ovu temu od 2007. godine do 2015. Faktori koji su označeni kao potencijalni prediktori su starost, nivo amputacije, funkcionalni status pri prijemu, kao i komorbiditeti. Međutim, u radu [9], koji se takođe bavio pregledom literature na ovu temu, jedino je uočena snažna veza između parametra balans i sposobnosti hoda uz pomoć proteze.

Poglavlje 2 sadrži opis dostupnog skupa podataka, kao i mehanizme za rešavanje problema nedostajućih podataka i balansiranje skupa podataka. U poglavlju 3 dat je opis predloženih metoda i izbora obeležja za klasifikaciju. Ostvareni rezultati i diskusija predstavljeni su u poglavlju 4. Naposletku, u zaključku, u sklopu poglavlja 5, dat je rezime i naznačeni su pravci daljeg isteaživanja.

II. Eksperimentalni podaci

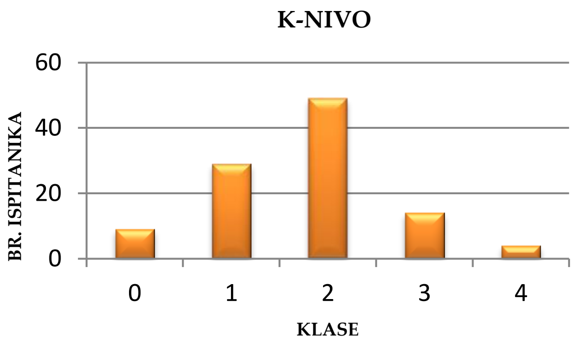
U radu je analiziran uticaj različitih atributa na predikciju ishoda rehabilitacionog tretmana pacijenata sa amputacijom donjih ekstremiteta. Podaci koji su korišćeni u istraživanju, prikupljani su u Kliničkom centru Vojvodine u period između 2010. i 2012. godine. Bazu podataka činila su 104 pacijenta, različitog pola i starosti (62,1 ± 10,9 godina) koja su bila podvrgnuta eksperimentu. Kriterijum za uključivanje pacijenata u istraživanje bila je jednostrana amputacija donjih ekstremiteta iznad nivoa skočnog zgloba pacijenata koji su prvi put snadbeveni protezom. Svi ispitanici dali su pismeni informativni pristanak za učešće u ovoj studiji. Istraživanje je dobilo saglasnost za sprovođenje od strane Etičke komisije Medicinskog fakulteta u Novom Sadu i Kliničkog centra Vojvodine.

A. Balansiranost skupa podataka

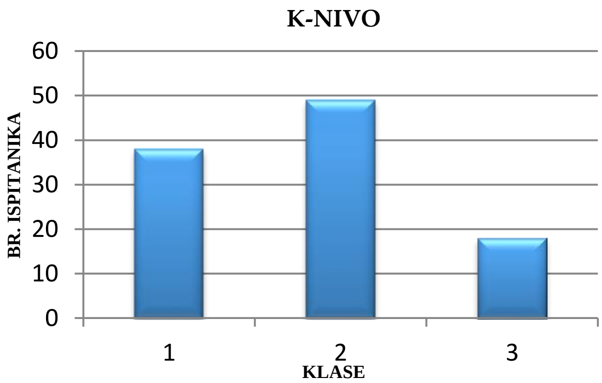
Kao što je već napomenuto, 104 pacijenta je učestvovalo u istraživanju, a njihova klasna raspodela za kategoriju K-nivo je prikazana na slici 1. Sa grafika se može očitati da je međuklasna razlika u broju ispitanika veoma velika.

Klasifikatori mašinskog učenja ne mogu se nositi sa nebalnsiranim setom podatka, odnosno ovi algoritmi imaju tedenciju da favorizuju klasu sa najvećim udelom ispitanika [10]. Ovakva neuravnoteženost može posebno da bude problematična kada nas zanima tačna klasifikacija manjinskih klasa, klasa sa najmanjim brojem ispitanika, kao što su u našem slučaju klase označene kao 0 i 4.

K-nivo predstavlja međunarodno priznatu skalu koja se koristi za predstavljanje ishoda rehabilitacionog tretmana pacijenata sa amputacijom donjih ekstremiteta. Donja granica K-nivoa, označena klasom 0, predstavlja nemogućnost samostalnog hoda uz pomoć proteze, dok je maksimalni ishod rehabilitacionog tretmana označen klasom 4. Sa slike 1 se očitava velika međuklasna razlika u broju ispitanika.



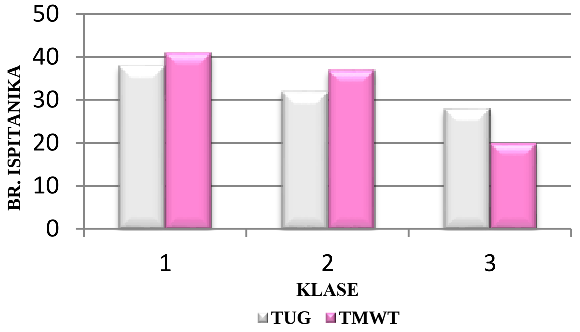
Sl. 1. Raspodela ispitanika prema klasama za kategoriju K-nivo.



Sl. 2. Modifikovane klase kategoriju K-nivo.

Da bi se smanjila neuravnoteženost broja ispitanika po klasama, izvršena je modifikacija klasa (slika 2), po uzoru na [3, 4]. Novi nivo 1 označavao bi nemogućnost samostalnog hoda uz pomoć proteze ili strogo ograničen hod na veoma kratkim relacijama (hod u kućnim uslovima), praktično, sama proteza ne bi značajno poboljšala mobilnost pacijenta, niti njegov kvalitet života. Posle modifikacije, ovoj klasi pripadalo je 38 ispitanika. Nivoom 2 predstavljao bi se ishod rehabilitacije pacijenata koji su imali mogućnost hoda na relacijama koje bi bile i izvan kuće ali uz značajna ograničenja, dok pacijenti nivoa 3 bi bili osposobljeni za hod na dugim relacijama, uz minimalna ili čak bez ograničenja. Nakon modifikacije, ovim grupama pripadalo je 49, odnosno 17 ispitanika, respektivno.

Na osnovu rezultata testa dvominutnog hoda i testa ustani i kreni, formirane su nove kategorije, koje bi potencijalno predstavljale ishode rehabilitacionog tretmana. Ukoliko bi za vreme testa dvominutnog hoda ispitanik prešao manje od 26 metara, klasifikovan bi bio u klasu TMWT1. U slučaju da bi ispitanik prešao između 25 metara i 55 metara, pripadao bi klasi TMWT2, a ukoliko bi prešao više od 55 metara za vreme dvominutnog testa pripadao bi klasi TMWT3. Što je veći indeks klase, odnosno što je ispitanik prešao više metara za vreme testa, ishod rehabilitacionog tretmana je bolji.



Sl. 3. Raspodela ispitanika za test ustani i ktreni - TUG (belo) i test dvominutnog hoda - TMWT (roze).

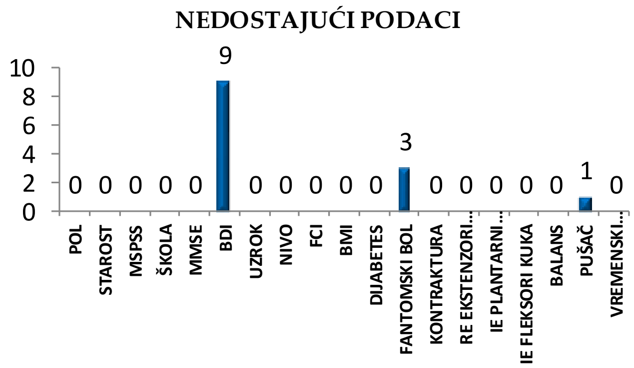
Test ustani i kreni 28 ispitanika završilo je za manje od 30 sekundi i svrstano je u klasu TUG3, 32 ispitanika klasifikovano je u TUG2, odnosno bilo im je neophodno između 30 i 60 sekundi za test, dok je njima 38 bilo potrebno više od 60 sekundi da bi završili test. Ovi podaci grafički su predstavljeni na slici 3.

Modifikacijom klasa smanjena je razlika broja ispitanika među klasama ali je i dalje postojala razlika koja bi potencijalno mogla uzrokovati favorizaciju većinske klase. Dodatno, ta razlika je smanjenja ponavljanjem uzoraka manjinske klase na slučajan način.

B. *Rešavanje problema nedostajućih podataka*

Istraživanje, koje je realizovano u Kliničkom centru Vojvodine, sprovedeno je kao prospektivna serija slučaja, što za posledicu ima mali broj nedostajućih podataka u bazi. Od 19 parametara kojima su predstavljeni pacijenti, kod samo 3 parametra se javljaju nedostajući podaci (slika 4).

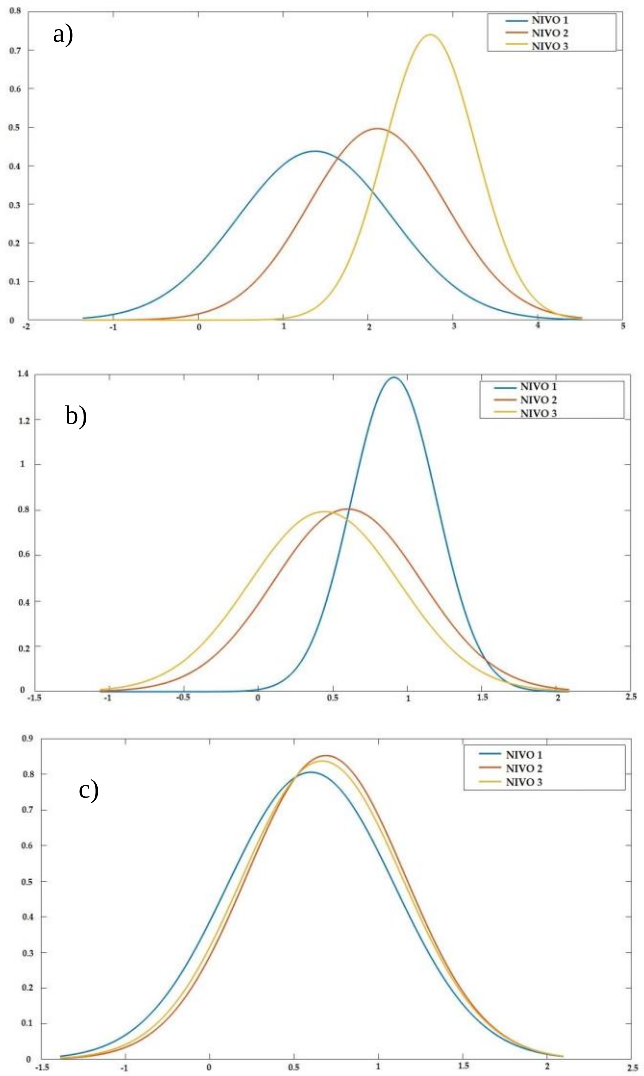
Jedan od najpopularnijih načina rešavanja problematike nedostajućih podataka jeste zamena nedostajućih podataka uzoračkom srednjom vrednošću ili modom [11]. Iako se ovom metodom umanjuje varijabilnost podatak (varijansa) i procene kovarijanse i korelacije u podacima (jer se ignoriše odnos između varijabli), zbog malog broja nedostajućih podataka smatrano je da ove promene nemaju statistički uticaj na konačni rezultat.



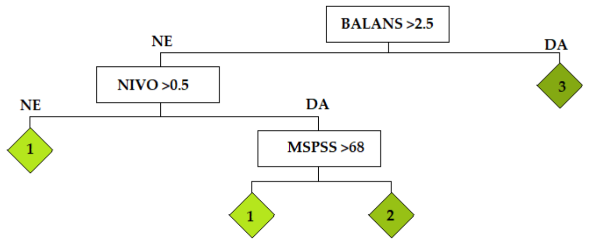
Sl. 4. Raspodela nedostajućih podataka prema parametrima. Nedostajući podaci se javljaju kod obeležja Bekova skala depresivnosti, fantomski bol i obeležja pušač.

III. Izbor obeležja i metode za klasifikaciju

Izbor adekvatnih obeležja ima ključan uticaj kako na kvalitet, tako i na efikasnost klasifikacije. Odabir (selekcija) obeležja podrazumeva biranje bitnijih (diskriminatornih) obeležja iz celog skupa podataka [11]. Na slici 5 prikazane su Gausove krive za obeležja balans, nivo amputacije i fantomski bol. Sa slike zaključujemo da obeležje fantomski bol nije diskriminatorno. Srednja vrednost obeležja svake klase, prikazane na apcisi, imaju slične vrendosti. Takođe, i verovatnoće ovog obeležja po klasama su veoma slične, pa je na osnovu ovog obeležja gotovo nemoguće klasno razdvojiti ispitanike. Međutim, obeležja balans i nivo amputacije moguće je okarakterisati kao diskriminatorna. Za potvrdu izbora obeležja i određivanje njihove značajnosti, korišćena je i tehnika stable odluke.



Sl. 5. Gausove krive za obeležje balans (a), nivo amputacije (b) i fantomski bol (c). Na osnovu ovih grafika možemo da zaključimo da su obeležja balans i nivo amputacije diskriminatorna, dok obeležje fantomski bol nije.



Sl. 6. Stablo odluke kako pokazatelj zančajnosti obeležja balans. Sa grafika je jasno uočljivo da veoma lako možeo izdvojiti klasu 3, odnosno pacijente sa potencijalnim visokim ishodom rehabilitacije.

A. *Stabla odluke*

Zbog svoje sistemske strukture, metoda stabla odluke je jednostvna i razumljiva za ljude. Izdvajanjem puta od odgovoarajuće klase, odnosno lista, pa sve do korena stabla, dobija se odgovor zašto je donesena neka odluka, pri čemu se iz svakog čvora čita razlog trenutnog izbora. Takođe, svako stablo odluke se jednoznačno može definisati preko skupa pravila "ako-onda" (engl. *if-then*), koja su osnovni gradivni blokovi baza znanja ekspertskih i drugih sistema zasnovanih na znanju [12].

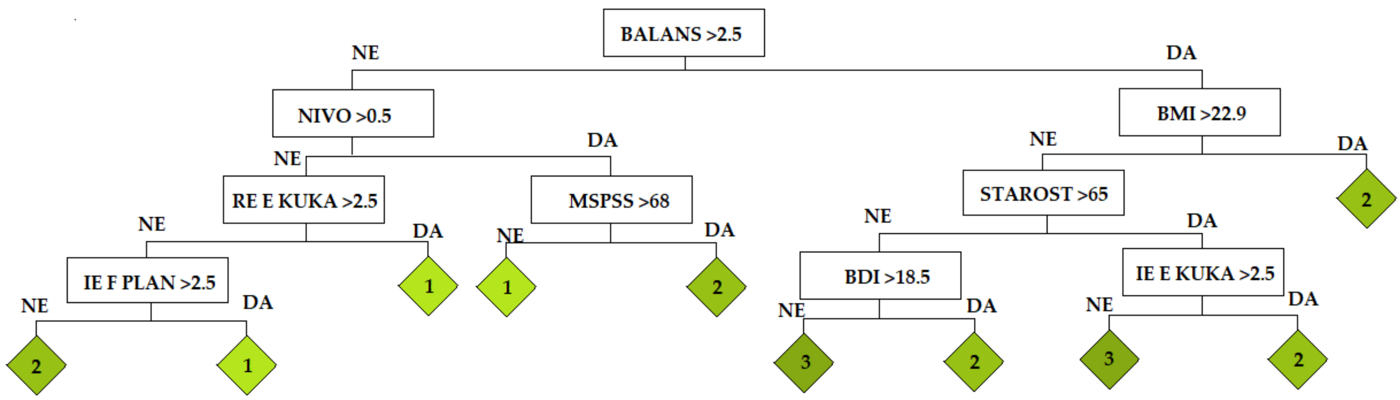
Stabla odluke izvršavaju klasifikaciju u dve ili više klasa, na osnovu vrednosti obeležja kojima opisujemo uzorke, propuštajući ih niz stablo od korena ka listovima. Na početku klasifikacije bira se obeležje čija vrednost najbolje deli raspoložive uzorke. Algoritam ID3, koji je razvijen za učenje stable odluke i koji je korišćen u ovom radu, obeležja deli na osnovu statističke veličine koja se naziva informacioni blok i koja se definiše preko entropije (entropija je 0 ako svi uzorci pripadaju istoj klasi a 1 ako klase imaju isti broj uzoraka) [12]. Analizom baze podataka, algoritam je procenio da je najbitnije obeležje balans. Sa slike 6 može se videti da se pomoću obeležja balans lako mogu izdvojiti pacijenti čiji bi ishod rehabilitacionog tretmana bio maksimalan (što je ujedno i potvrda za za Gausovu krivu sa slike 5).

Na slici 7 prikazano je stablo odluke sa izdvojenim parametrima za kategoriju K-nivo. Obeležja koja je stablo označilo kao diskriminatorna su: balans, nivo amputacije, starost, indeks telesne mase (engl. *body mass index* - BMI), mišićna snaga ekstenzora kuka rezidualnog ekstremiteta (*RE E kuka* na grafiku), Bekova skala depresivnosti (engl. *Beck depression inventory* - BDI), starost pacijenta pri prijemu, multidimenzionalna skala ostvarene socijalne podrške (engl. *multidimensional scale of perceived social support test* - MSPSS), kao i parametre mišićne snage ekstenzora kuka (*IE E kuka* na grafiku) i plantarnog fleksora (*IE F plan* na grafiku) intaknog ekstremiteta.

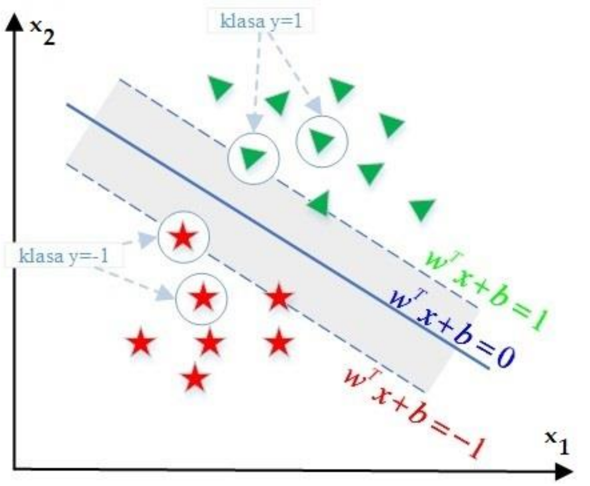
B. *Metoda vektora nosača*

Odabrana obeležja prosleđuju se algoritmima baziranim na vektorima nosačima. Primarna verzija SVM algoritma na ulazu uzima skup podataka i zatim određuje kojoj od dve moguće klase svaki uzorak pripada, odnosno traži se funkcija $f(x)$ koja je u funkciji hiper-ravni H, koja predstavlja razdvajajuću marginu za dva stanja sistema u prostoru obeležja [11], $H: \mathbf{w}^T\mathbf{x} + b = 0$, kao i dve hiper-ravni (koje se nazivaju vektori nosači) $H_1: \mathbf{w}^T\mathbf{x} + b = 1$ i $H_2: \mathbf{w}^T\mathbf{x} + b = -1$, uz uslov da ne postoje tačke između H_1 i H_2 i da je razmak između margina maksimalan. Šematski, to je predstavljeno na slici 8.

Iako je SVM algoritam primarno razvijen za binarnu klasifikaciju, on se može koristiti i za problem višeklasne klasifikacije. Standardan algoritam za rešavanje ovakvog problema je rastavljanje problema od M klasa na seriju problema od po dve klase, i konstrukcijom više binarnih klasifikatora [11]. Jedan od pristupa je metod „jedan protiv jedan“, koji formira za svaki par klasa po jedan klasifikator [12]. Ti klasifikatori su osposobljeni za razlikovanje uzoraka jedne klase od uzoraka druge klase. Klasifikacija nepoznatog uzorka se vrši prema maksimalnom broju „glasova“, gde svaki klasifikator „glasa“, za jednu klasu.



Sl. 7. Stablo odluke za ketegoriju K-nivo.



Sl. 8. Hiper-ravan *H* i njene margine u ravni parametara x_1 i x_2. Zelenim trouglovima predstavljeni su uzorci koji pripadaju klasi y=1, dok su zvezdicama predstavljeni uzorci klase y=-1.

IV. Rezultati i diskusija

Za generisanje i simulaciju rada SVM predikcionog modela i stabla odluke korišćen je programski paket *Matlab*. Pripremljeni podaci prosleđeni su algoritmu stabla odluke, koji radi po principu ID3 algoritma. Obeležja izabrana na ovaj način prosleđena su SVM algoritmu. Za rešavanje ovog problema izabran je linearni kernel. Rezultati su dobijeni evaluacijom obučenih modela nad odgovarajućim test skupovima, dobijenih krosvalidacijonom metodom sa 10 particija. U svakoj iteraciji koristilo se 9 različitih particija za obuku, odnosno 9 podskupova skupa svih uzoraka, dok se preostali skup koristio za testiranje.

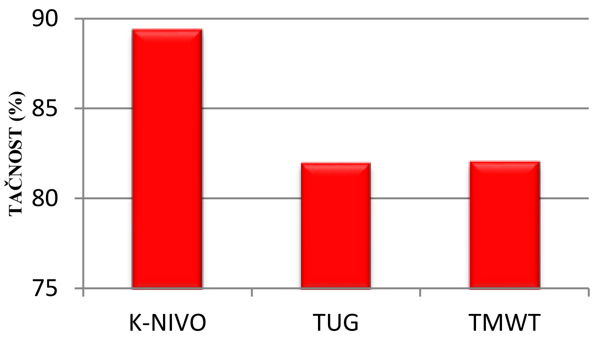
Tabela I predstavlja matricu konfuzije SVM klasifikatora za K-nivo, nakon rešavanja problema nedostajućih podataka i balansiranja klasa. Iz tabele se može iščitati da algoritam u veoma visokom procesu tačno klasifikuje. Tačnost SVM klasifikatora, odnosno procenat korektnih predviđanja od ukupnog broja predviđanja, iznosila je 89.3%. Tačnost klasifikacije ostalih kategorija prikazana je na slici 9. Najveća tačnost postignuta je prilikom predviđanja kategorije K-nivo (89,3%), dok najlošije performanse se javljaju kod klasifikacije kategorija testa ustani i kreni (81,88%). Pored tačnosti, i ostale mere za evaluaciju klasifikatora za kategoriju K-nivo su veoma visoke. Osetljivost klasifikacije, odnosno procenat korektno predviđenih od ukupnog broja pravih ponavljanja te klase za K-nivo iznosi 83,95%, specifičnost (sposobnost testa da identifikuje odsustvo neke klsse) 91,97%, dok je preciznost (odnos tačnih predviđanja i ukupnog broja predviđanja) iznosila 83,7%.

V. Zaključak

Uzimajući u obzir stalno povećanje broja pacijenata sa amputacijom donjih ekstremiteta, kao i cenu i trajanje rehabilitacionog tretmana, predikcija ishoda rehabilitacionog tretmana predstavlja jedno od veoma aktuelnih pitanja savremene rehabilitacione medicine. Brojni radovi napisani na

TABELA I
MATRICA KONFUZIJE ZA KATEGORIJU K-NIVO

		Predviđene klase		
		1	2	3
Prave klase	1	**50**	4	0
	2	7	**38**	9
	3	0	6	**48**



Sl. 9. Tačnost klasifikacije za različite kategorije.

ovu temu, čiji je jedan mali deo korišćen kao smernice i inspiracija ovom radu, svedoče o značaju ove tematike.

Glavni cilj razvijanja sistema za predikciju uspešnosti rehabilitacionog tretmana treba da bude identifikacija tačno onih obeležja koja će doprineti uspešnoj predikciji, to jest pronalasku onih parametara koji će omogućiti korektnu identifikaciju pacijenata sa amputacijom pogodnih za propisivanje proteze i uključivanje u odgovarajući rehabilitacioni tretman. Obradom podataka i selekcijom obeležja metodom stable odluke, postignute su veoma visoke performance SVM klasifikatora. Obeležje koje je označeno kao balans izdvojeno je kao jedan od najznačajnijih prilikom identifikacije pacijenata pogodnih za propisivanje proteze. Maksimalna efikasnost postignuta je prilikom predikcije K-nivoa i iznosila je 89.30%.

Ovakav sistem veštačke inteligencije, zbog svojih visokih performansi, moguće je primeniti i u kliničkoj praksi. Dalji rad u budućnosti trebalo bi da se skoncentriše na testiranju algoritma nad proširenom bazom podataka, odnosno predviđanje nad originalnom petostepenom skalom.

Zahvalnica

Ovaj rad podržan je od strane Ministarstva prosvete, nauke i tehnološkog razvoja kroz projekat broj 451-03-68/2020-14/200156: “Inovativna naučna i umetnička ispitivanja iz domena delatnosti FTN-a”.

LITERATURA

[1] A. Knežević, “Faktori koji utiču na nivo osposobljenosti za hod uz pomoć proteze nakon amputacije donjih ekstremiteta,”Doktorska disertaciija, Katedra za fizikalnu medicine i rehabilitaciju, Medicinski fakultet, Univerzitet u Novom Sadu, Srbija, 2014.

[2] K. K. Chui, M. M. Jorge, S. C. Yeng, M. M. Lusardi, *Orthotics and Prosthetics in Rehabilitation, fourth edition,* St. Louis, Missouri, Elsevier, 2021.

[3] A. Knežević, M. Petković, A. Mikov, M. Jeremić Knežević, Č. Demeši Drljan, K. Bošković, S. Tomašević Todorović, Z. Jeličić, “Factors that predict walking ability with a prosthesis in lower limb amputees,” -Srpski arhiv za celokupno lekastvo, 507-513, Sept-Okt 2016.

[4] N. Majdič, G. Vidmar, H.Burger, “Establishing K-levels and prescribing transtibial prostheses usig six-minute walk test and one-leg standing test on prosthesis: a retrospective audit,” International Journal of Rehabilitation Research, 266-271, Sep 2020.

[5] B. Majstorović, M. Pešta, “Factors predicting rehabilitation outcome in patiens after unilateral transtibial amputation due to peripheral vascular disaese,” Military Medical and Pharmaceutical Journal of Serbia, 357-362, Maj 2020.

[6] T. Pohjolainen, H. Alaranta, “Predictive factors of functional ability after lower limb amputation, ” Annales chirurgiae et gynaecologice, 36-9, 1991.

[7] K. Sansam,V. Neumann, R. O’Connor, B. Bhakta, “Predicting walking ability following lower limb amputation: a systematic review of the literature,” Journal of Rehabilitation Medicine, 593-603, Jul 2009.

[8] J. T. Kahle, M. J. Highsmith, H. Schaepper, A. Johannesson, M. S. Orendurff, K. Kaufman, “Predicting walking ability following lower limb amputation: an updated systematic literature review,” Technology and Innovation, 125-137, Sep 2016.

[9] J. Van Velzen, C. van Bennekom, L. van der Woude, H. Houdijk, “Physical capacity and walking ability after lower limb amputation: a systematic review,” Journal of Clinical Rehabilitation, 999-1016, Nov 2016.

[10] C. Micheloni, A. Rani, S. Kumar, G. L. Foresti, “A balanced neural tree for pattern classification,” Elsevier: Neural Networks, 81-90, Mar 2012.

[11] V. Crnojević, “Prepoznavanje oblika za inženjere,” Novi Sad, Srbija, Fakultet tehničkih nauka u Novom Sadu, edicija: Tehničke nauke, 2014.

[12] M. Milosavljević, “Veštačka iinteligencija,“ Beograd, Srbija, Univerzitet Singidunum, 2019.

ABSTRACT

Prosthetic rehabilitation is currently the best treatment for patients with lower limb amputation. However, prosthesis fabrication and accompanying prosthetic rehabilitation are very long and expensive process that sometimes does not lead to improved mobility and quality of life of patients. Therefore, it is necessary to predict the result of rehabilitation treatment. The main task of this work was to make a tool, based on artificial intelligence algorithms, which could be applied as early as possible, to make the best possible prediction of the result of rehabilitation of patients with amputation of the lower extremities, or to predict the K-level (Medicare Functional Classification Level), Walking Ability Level, Two-Minute Walk Test and Timed Up and Go Test. Evaluation of the model performed on data from the Rehabilitation Clinic, Clinical Center of Vojvodina. The obtained results show a significant improvement, in terms of classifier performance, compared to previous methods and confirm the choice of some of the most important parameters in patient identification.

Predicting results of prosthetic rehabilitation in lower limb amputees by using artificial intelligence algorithms

Jovana Arsenović, Aleksandar Knežević, Mirna N. Kapetina and Zoran D. Jeličić

Multipurpose remote monitoring system based on microservice architecture

Luka Bjelica, Miloš Panić, Marko Pejić

***Abstract*—This paper presents a cloud-native, multi-purpose, and reusable system for collecting, processing and storing data, with the aim of monitoring an arbitrary physical system. The proposed system can be divided into three main parts: a private network containing a set of microservices that perform complete data processing, applications that implement the low-level logic for collecting data from remote sensors, and a web client which enables interaction between the user and the rest of the system. The final product of this paper is a system based on the microservice architecture named *isobar.ot*, that allows monitoring of the chosen set of values of an arbitrary physical system, through a simple and functional user interface. Using the system presented in this paper, the user is able to control the entire course of remote monitoring: from the selection and specification of the collected data scheme, through the definition of alarm values, to displaying changes of values and alarms in real-time.**

Keywords: distributed systems, microservice architecture, remote monitoring systems, cloud-native systems, Internet of Things

I. Introduction

Monitoring the various parameters of arbitrary physical systems is a crucial part of every industrial facility. Supervisory Control and Data Acquisition (SCADA) systems are ubiquitous in almost all industries: from the food industry to the power industry, which results in a need for continuous improvements of the existing, and development of new solutions [1].

This paper concerns the development of the modern solution for remote monitoring systems that can be used for monitoring an arbitrary physical system. It is based on a microservice architecture with cutting-edge tools and technologies. Motivation for choosing this topic came from a necessity for a system that can work with large amounts of data and is flexible in a relation to a supervised physical system, which makes it usable as a part of Internet of Things (IoT) systems [2].

One of the biggest challenges when designing such a system is the scalability, i.e., the ability of the system to work with a large number of sensors and serve a large number of clients without a drop in performance. Furthermore, such a system requires a simple and functional user interface in order to provide an operator with an efficient way to monitor changes in the collected data, have insight into the alarming events in real-time, as well as defining new locations, alarm types, schemes of the data that is collected, etc. For the above-mentioned goals to be fulfilled, the proposed system is designed according to the principles of microservice architecture.

After the Introduction, basic principles, advantages and disadvantages of the microservice architecture and the architecture of the proposed solution are explained in Section II. Technical details about the implementation, along with the tools and technologies that were used are introduced in Section III. Results and user interface are shown in Section IV and the concluding remarks along with future plans are given in the final Section V.

II. Architecture

Two mandatory requirements that the proposed system must meet are working with a large amount of data and serving a large number of clients. The architecture of the proposed system is designed so that the mentioned requirements are satisfied for the arbitrary amount of data and number of clients.

Microservice Architecture

Microservice architecture implies the development of applications in the form of small, isolated, and independent services that communicate with each other via clearly defined protocols. Such a method for developing systems came about due to the aim of overcoming flaws and problems that come with monolith architecture.

The traditional approach to developing software implies the use of monolithic architecture. One of the flaws of monolithic architecture is that it poorly copes with overload [3]. One approach to handling overload is to vertically scale the existing machine on which the application runs. That is expensive and not efficient enough to solve the problem completely. The other approach is to use horizontal scaling and create multiple instances of the application. This approach leads to inefficient use of hardware resources because there is no possibility to scale only those parts of the system that require it [3]. In addition, the degree of reusability of individual components is reduced because they are tightly coupled with the system they were initially developed for.

The main advantage of the microservice architecture-based systems is that the individual microservices can be scaled independently. That way, better utilization of hardware resources is achieved so that only the parts of the system that are affected by overload get scaled. Additional advantages that the microservice architecture brings are independent

L. Bjelica (bjelicaluka@uns.ac.rs), M. Panić (panic.sw19.2018@uns.ac.rs), M. Pejić (markopejic@uns.ac.rs), University of Novi Sad, Faculty of Technical Sciences, Department of Computing and Control Engineering, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia.

development of individual components as well as the high degree of their reusability [3].

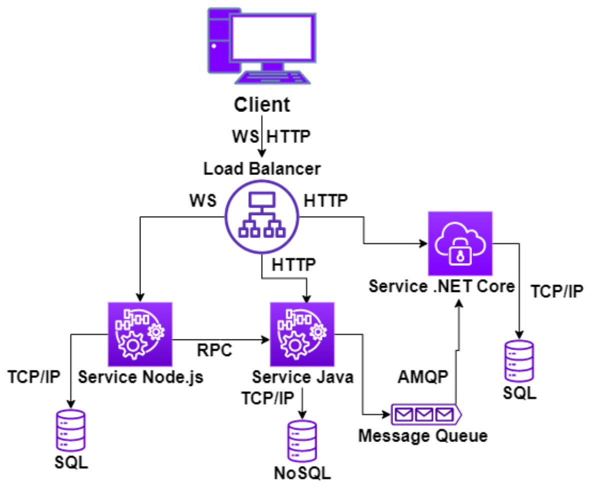


Fig. 1. Microservice Architecture Scheme

Fig. 1 presents the example of the microservice architecture scheme. It can be noticed that system based on the microservice architecture is suitable for scaling individual parts that require it because each part represents an independent application. There is also the possibility of using different technologies for implementing individual services, as well as the use of different communication mechanisms and protocols depending on the need.

System development follows the *divide and conquer* principle, which divides a large and complex system into smaller units, which are easier to develop. It is important to mention that dividing the system into smaller entities does not solve the complexity problem, but rather it delegates it to a level above, that is, to connect the system's components and their orchestration.

Previously stated benefits surpass hardware limitations of a single machine on which the system is running and thus make the microservice architecture an adequate solution for implementing remote monitoring systems.

System Architecture

The proposed system represents a set of components, each being an independent application with a unique role. The system is made out of services that are responsible for: authentication, user groups and user profiles, schemes of data that is collected from arbitrary physical sensors, validation, persistence and aggregation of the collected data, detecting alarms, generating reports from the aggregated data, and displaying the user interface.

The architecture of the proposed system is presented in Fig. 2. The system runs on the cloud and is completely independent of the physical system from which it receives the data.

The responsibility for collecting data from physical sensors and sending it to the system is encapsulated within the Local Processing Unit (LPU). LPU acts as an intermediary

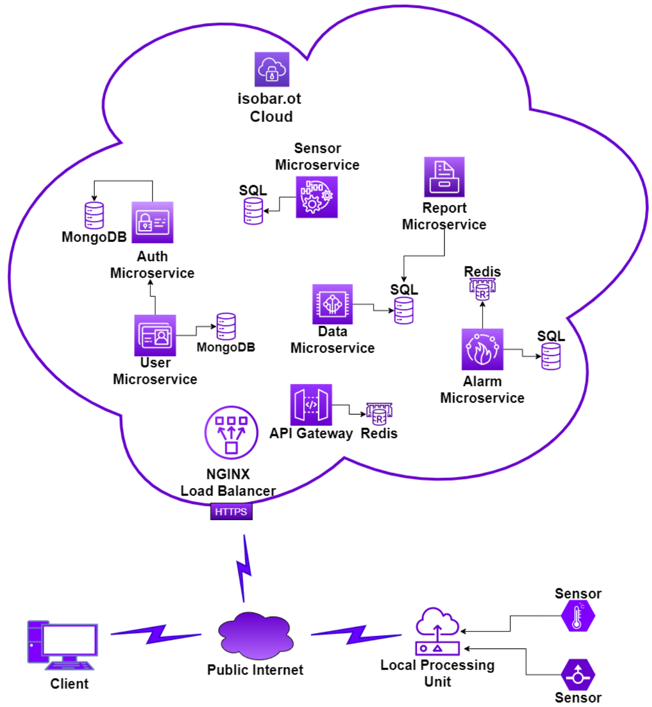


Fig. 2. System Architecture Scheme

between physical sensors which collect raw data, and the part of the system which is launched on a cloud.

The services are a part of the private, isolated network and it's not possible to address them directly from the public internet. The ingress service, acting both as a reverse proxy and load balancer [4], is the only way to access services from the outside. All inter-service communication is done in a synchronous manner with one exception. The only asynchronous communication is data ingress where the API Gateway service sends the data to the Data and Alarm services separately and is not interested in the content of the response.

A brief description of individual services that make up the system follows:

1) Authentication Service: The role of the authentication service is to provide a safe way to persist user credentials. This service implements the logic for generating access tokens that are used by other services for restricting unauthorized access to individual resources. It's worth pointing out that the authorization is highly dependent on the context it's used in and implies that details of role-based access control (RBAC) used for restricting access to individual resources need to be defined in the service that owns the resources. Otherwise, every service that implements RBAC would be tightly coupled to the service that implements authorization logic. The problem with that approach is that the authorization service may represent a bottleneck of the system. Tight coupling of services is not a problem if they are running inside the same process. Knowing that the presented system is distributed, this approach presents a big problem because it increases the degree of inter-service communication and causes the known problem "chatting" [5]. The problem mentioned above is the reason why the authorization is not implemented within the authentication service.

2) User Service: The service responsible for working with user profiles and user groups is a very important component of the system because it enables isolating data at the level of user groups. This service is responsible for protecting user's personal information and controlling which user group does the user belong to. Considering that most of the resources on the system are tied to a specific user group, user has access limited only to those resources tied to the group it belongs to.

3) Sensor Abstractions Service: A key component of the proposed system is a service that provides a way for arbitrary sensors that collect data on a supervised physical system. The role of this service is to persist information about the schemes of data that is being collected, as well as information about the concrete sensors that send that data. The entities used by this service are sensor abstractions, i.e., schemes of data that is being collected, and the information about the concrete physical sensors along with their locations. The idea behind defining data schemes is the possibility of validation of incoming data on the service, as well as the possibility of using the same data scheme for multiple different sensors. The scheme represents a set of individual tags (physical values of interest), each containing a name and a primitive data type. Apart from the name and the simple type, for every tag in the data type, aggregation methods are listed, based on which, aggregation service knows how to process raw data.

4) Data Validation Service: Received data first goes through a validation process, which is carried out based on the previously defined data type. This service is also responsible for verifying the validity of the public API token, using which the LPU unit proves authenticity. Apart from validation, this component presents a suitable place for dispatching events about received data in real-time. Events are dispatched through previously defined bidirectional communication protocol with the aim of achieving *publisher-subscriber* mechanism [6].

5) Data Persistence and Aggregation Service: The component which contains markedly the most complex logic and which requires the most hardware resources is data persistence and aggregation service. Before it gets aggregated, the raw data are persisted inside a temporary data store that is being cleared after a fixed period of time. The reason why raw data aren't stored permanently is that the amount of data is immensely large and that storing it isn't efficient. The more efficient solution is doing periodical data aggregation, such that users are able to define a time period after which the aggregation is performed. Additionally, users are able to define the methods by which data aggregation is performed, which later allows them to follow trends and generate reports of interest. By that, the system gets better performances, not only in terms of memory usage but also in decreasing the time needed for generating certain reports.

6) Report Service: The purpose of the persisted data lies in the ability to generate certain reports from them, with the aim of monitoring trends and presenting behavior of arbitrary values that are collected. This service implements the logic for generating reports on the aggregated data that is permanently persisted in the system. The report takes into account the specific frequency at which the data was aggregated as well as the time interval within which the data was collected. It also provides the ability to define and store report types that contain all the information needed to generate a particular report, except for the time interval.

7) Alarm Service: Detecting critical values, that is, data values which deviate from predefined boundaries can be very significant for physical systems which the proposed system is monitoring. The responsibility of this component is the detection of critical values and dispatching events about them, in real-time. Critical values are detected by rules previously defined in alarm types. Alarm type contains priority, a threshold value, and the information about whether the threshold presents an upper or lower limit of the normal state. A property from the data type can have a set of predefined alarm types tied to it. During alarm detection, every alarm type that is tied to a certain property is taken into consideration. When the critical value is detected, an event is dispatched through a predefined, real-time communication protocol. After the alarm event is dispatched, the client has the option of caching that alarm for a certain time period and thus preventing the system from dispatching more of the same events tied to the alarm of a certain priority, type, and limit value.

8) User Interface Service: This service provides elements needed for the graphical presentation of real-time data and generated reports. In addition to that, it contains elements that can be used to create certain resources, set certain rules, and take care of users and user groups.

III. Implementation

The microservice architecture allows the usage of numerous technologies for implementing individual components so that the most suitable technology for the requirements specific to that component is used. Fig. 3 shows an overview of all technologies used for implementing certain parts of the system.

Implementation of Individual Components

The authentication service implementation was realized using the .NET Core [7], while the MongoDB [8] database was used for the persistence of user accounts. Each user account consists of a unique name, password, and role. In case of the data leak, *hashing and salting* [9] of passwords is applied with the aim of preventing their misuse.

For the purpose of implementation of the service for working with user groups and profiles, .NET Core was used. User groups and profiles are in a one-to-many relationship, i.e., a profile belongs to exactly one group, while a group can contain several user profiles. The user group contains a name, surname, and e-mail address. To ensure that the

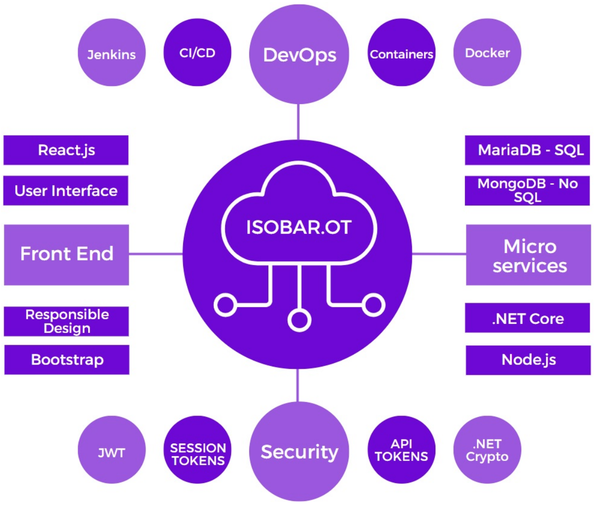


Fig. 3. Overview Of The Used Technologies

connection between user groups and accounts is modeled properly, MariaDB [10] relational database was used.

Another of the services implemented using .NET Core is a service for defining abstractions of physical sensors, i.e., schemes of data that come from remote sensors. Entities relevant to this microservice are sensor schemes, physical sensors, and locations. Since these entities are interrelated, a relational database MariaDB was used for data persistence. Primitive data types which are supported are real numbers, boolean values, enumerations, and text. Each sensor has its own public API token, which is a randomly created string of letters and digits that can be altered. Altering the public API token of a sensor is a significant function since it provides a mechanism of protection from receiving false data from a party that has obtained the token in an unauthorized way.

The implementation of the data validation logic is separated into the data validation service implemented in the Node.js [11] environment. This service uses Redis [12] cache as temporary storage for a public API key as well as the data scheme of the authenticated sensor. This decreases communication with the service in charge of validating public API tokens and evades creating the system's bottleneck. Sending data in real-time is done by using a WebSocket, using the socket.io library [13].

Data persistence and aggregation service is implemented using .NET Core and MariaDB database. One of the main reasons for choosing MariaDB as a relational database is the native support for built-in mechanisms that can be used for storing and manipulating data in dynamic JSON [14] format. The collected data is stored in a temporary table whose content is deleted after an all-day cycle of aggregations. The system supports several different aggregation time periods, of which the smallest is five minutes, and several different methods for aggregating real numbers including minimum, maximum, sum and mean. The aggregated data is stored in separate tables in the database, each corresponding to a single resolution.

.NET Core was used for the implementation of the report service. This service reads data from the database in which the persistence and data aggregation service has stored the processed data. Report generation was realized with the help of mechanisms for manipulating data in JSON format that are supported by the MariaDB database.

Implementation of alarm service is done inside the Node.js environment, while the MariaDB database was used for storing information about the alarm schemes and concrete critical values themselves. Sending data in real-time is done using WebSocket, which makes users promptly informed about every critical value of the monitored system. This service also uses Redis cache for storing critical values to avoid notifying the client unnecessarily. Another role of Redis is to synchronize socket.io events between multiple instances of the application.

The user interface was implemented using React.js [15] and Bootstrap [16] libraries.

Automation of Development Processes

Developing a system that is based on a microservice architecture increases the maintenance complexity because the source code of the system is made up of multiple smaller and often independently maintained code bases. Continuous Integration (CI) represents a necessary part of the development process of systems composed of many components with independently maintained code bases. That includes validation and testing of individual functionalities, as well as the rebuilding of components affected by changes. Continuous Deployment (CD) is the process of automatic reflection of changes to the final system which is used in production. In systems that are subject to frequent changes, the CD represents a necessity and can be very important in both the development and deployment phases of the system. In the development process, an instance of the staging application is created in order to provide access to the application to everyone that is involved. That significantly increases the degree of error detection in the development phase and reduces the chances of a bug in production.

Developing the system comprised of components that are implemented in different technologies, complicates the requirements for the environment in which individual components can be started. The concept of containers is introduced with the aim of providing a virtual environment at the operating system level, which can be predefined, packaged, and quickly launched. Such an environment has a high degree of portability and can be run on any platform on which a container engine can be run. The proposed system uses Docker [17] for the containerization of individual components. Docker provides an API for defining, packaging, transferring, and running virtual environments in form of containers. It also supports the creation of isolated private networks within which containers can intercommunicate and reference eachother using the local DNS server [18].

High availability and fault tolerance deserve special attention for distributed systems that are running in a production environment. A highly available system tends to minimize service downtime while a fault-tolerant one ensures that no data is lost during minor or major failures. The goal is to have both a highly available and fault-tolerant system that ensures high service uptime without data loss. The fact that the proposed system can be deployed on a multi-node cluster and that the data replication on multiple nodes is supported ensures that the aforementioned goals are met in a production environment. By distributing traffic across multiple nodes, the uptime of the individual services, resilience to data loss, and the request processing capacity are increased.

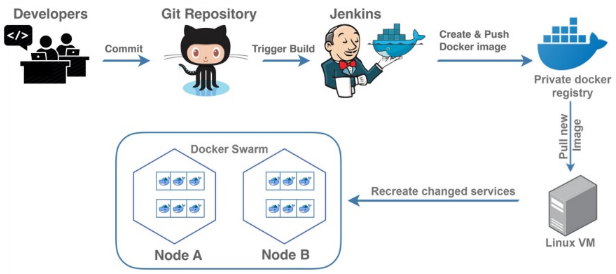


Fig. 4. Automatic Development Process Scheme

Fig. 4 illustrates the scheme of the automated CI/CD pipeline used by the proposed system in a development environment. Once the changes are made, they are synchronized with the remote codebase located on GitHub. The application that implements the CI/CD pipeline, i.e. Jenkins [19] gets notified, via the WebHook submission mechanism, about the changes that have been made on a certain branch. When an event containing information about the changes reaches Jenkins, a predefined pipeline is started. The pipeline runs the process of validating changes and rebuilding Docker Images if the changes prove to be valid. Additionally, the affected parts of the system get synchronized with newly integrated changes by a command that gets executed using SSH (Secure Shell) on the machine on which the system is running.

The proposed system can be instantiated using the *docker-compose* tool, for the needs of the development environment or using the *kubectl* tool for production environments. The role of the aforementioned tools is to take care of pulling, configuring, starting, and shutting down previously defined services, creating private networks within which the services communicate and scaling certain services depending on the needs of the system.

IV. Results

The final product of this paper is a functional remote monitoring system, based on microservice architecture and modern technologies, called *isobar.ot*. The system described in this paper supports the processing and persistence of arbitrary data sets, as well as tracking of trends in the collected data, i.e., periodical changes in the data values and the detection of critical values in real-time. In addition, the system allows the generation of reports from the persisted data for a certain period of time, which allows the user to analyze the behavior of the physical system that is monitored.

The most valuable aspects of the proposed system are the fact that it can be reused for many different physical systems, and also its scalability which is ensured by the microservice architecture. The system solves one of the basic problems that come with the IoT systems, which is giving semantics to the raw data collected from the physical sensors so that storing and processing of data is realized in a uniform way, independent of the nature of data.

Besides various simulations, such as collecting data on weather conditions, that were used to test the system's reusability, there are two successful use-cases of the proposed system. Both of them are Road Traffic Monitoring Systems (RTMS) that are deployed in India and Croatia. The first receives data from two different sensor types: laser and radar. The laser sensor sends detailed information about the passed vehicles such as speed, class, and it's dimensions. The radar detects vehicle's speed and relative position and send them in real-time.The second use-case receives data from a camera that detects which class of vehicle has passed. The previously mentioned use-cases prove that the proposed solution can be used to monitor arbitrary physical systems. In both cases, the system ensures high availability and fault tolerance, i.e. it is deployed on a multi-node cluster and uses data replication in order to prevent data loss.

User Interface

Fig. 5 shows the appearance of the user interface, which monitors critical values in real-time. The hide option gives users an opportunity to declare that they are aware of a particular alarm, to take all necessary steps, and not want the system to notify them more about the alarm so that they can pay attention to other alarms.

PROPERTY	TYPE	CRITICAL VALUE	PRIORITY	LOCATION	SENSOR	
sinus	Above	79.1056663295	Low	Zrenjanin	Sinus Simulation	Hide
sinus	Above	86.9963516663	Medium	Zrenjanin	Sinus Simulation	Hide
sinus	Above	95.2598516295	High	Zrenjanin	Sinus Simulation	Hide
tanh	Below	-50.1444326271	High	Čurug	Tanh Simulation	Hide
cosinus	Below	-5.88190451025	Medium	Novi Sad	Cosinus Simulation	Hide

Fig. 5. Alarms Monitoring in Real-Time

The *Dashboard* section (Fig. 6) presents the appearance of the user interface through which data arrived from sensors is tracked. Selecting the sensors is done from the drop-down menu, where all the sensors belonging to the corresponding user groups are listed. It is possible to choose more than one sensor and to display data in real-time graphically and in a table.

For the purpose of testing the proposed system, simulation sensors are implemented, which represent individual applications independent of the rest of the system. Their role is to simulate the operation of LPU units by sending random values or values of certain mathematical functions, instead of collecting data from physical sensors. Some of the functions supported by simulation sensors are sine, cosine, sigmoid, ReLU (Rectified Linear Unit), and the like.

Fig. 6. Live Data Display

V. Conclusion

This paper presents a cloud-native, multi-purpose remote monitoring system, based on microservice architecture. The implemented system is named *isobar.ot* and consists of three main parts: local processing units (LPU) used for implementing low-level logic for collecting data from physical sensors and sending it to the cloud, an isolated private network with a set of microservices that perform the entire processing and persisting data on the cloud, a client web application that allows users to interact with the rest of the system.

The biggest advantage of the proposed system lies in its ability to monitor arbitrary physical systems, and the ability to work with a large number of sensors and serve a large number of clients. These advantages are achieved by relying on microservice architecture and modern technologies.

Despite the fact that inter-service communication is brought to a minimum, there are cases where a better solution for communication would be to use asynchronous protocols, such as AMQP [20]. Another disadvantage of the proposed system is that some services use data that are not owned by them, i.e., they have to turn to services that have that data. This reduces the failure resistance of interdependent parts of the system. A potential solution to this problem is to use caching more frequently or replicate the data used by multiple services while maintaining asynchronous consistency.

In addition to overcoming the previously mentioned shortcomings of the proposed system, the plan for further development is the implementation of a data export service in the form of a file of a certain format, such as PDF, CSV, JSON, and the like. Also, the plan is to implement support for receiving data in several different protocols, and not only in HTTP. Another possibility for further development of the proposed system is the implementation of a uniform control component. The task of this component is to provide a mechanism for managing the monitored physical system, at a high level. That way, the proposed system would be extended to a fully functional supervisory and control system, which is certainly in the plan for future development.

References

[1] A. Goel and R. Mishra, “Remote data acquisition using wireless - scada system,” *International Journal of Engineering*, vol. 3, 03 2009.
[2] L. Atzori, A. Iera, and G. Morabito, “The internet of things: A survey,” *Computer Networks*, vol. 54, no. 15, pp. 2787–2805, 2010.
[3] O. Al-Debagy and P. Martinek, “A comparative review of microservices and monolithic architectures,” in *2018 IEEE 18th International Symposium on Computational Intelligence and Informatics (CINTI)*, pp. 149–154, 2018.
[4] “Nginx documentation.” https://docs.nginx.com/nginx/admin-guide/load-balancer/http-load-balancer/, 2021.
[5] J. Ghofrani and D. Lübke, “Challenges of microservices architecture: A survey on the state of the practice,” 05 2018.
[6] C. R. Ozansoy, A. Zayegh, and A. Kalam, “The real-time publisher/subscriber communication model for distributed substation systems,” *IEEE Transactions on Power Delivery*, vol. 22, no. 3, pp. 1411–1423, 2007.
[7] “.net core documentation.” https://docs.microsoft.com/en-us/dotnet/, 2021.
[8] “Mongodb documentation.” https://docs.mongodb.com/, 2021.
[9] “Adding salt to hashing: A better way to store passwords.” https://auth0.com/blog/adding-salt-to-hashing-a-better-way-to-store-passwords/, 2021.
[10] “Mariadb documentation.” https://mariadb.com/kb/en/documentation/, 2021.
[11] “Node.js documentation.” https://nodejs.org/en/docs/, 2021.
[12] “Redis documentation.” https://redis.io/documentation, 2021.
[13] “Socket.io documentation.” https://socket.io/docs/v3/index.html, 2021.
[14] “Json documentation.” https://www.json.org/json-en.html, 2021.
[15] “React documentation.” https://reactjs.org/, 2021.
[16] “Bootstrap documentation.” https://getbootstrap.com/docs/5.0/getting-started/introduction/, 2021.
[17] “Docker documentation.” https://docs.docker.com/, 2021.
[18] D. Jaramillo, D. V. Nguyen, and R. Smart, “Leveraging microservices architecture by using docker technology,” in *SoutheastCon 2016*, pp. 1–5, 2016.
[19] “Jenkins documentation.” https://www.jenkins.io/doc/, 2021.
[20] “Amqp documentation.” https://www.amqp.org/, 2021.

Integrated Particle Filter for Multi Target Tracking

Zvonko Radosavljević, Dejan Ivković and Branko Kovacević

***Abstract*— Target tracking in heavy cluttered environment requires methodology for false track discrimination and data association. Recently, we present a new particle filter (PF) approach which recursively calculates the probability of target existence for the false track discrimination. Our approach treats possible detections of targets followed by other tracks as additional clutter measurements. It starts by approximating the a priori probabilities of measurement origin. The posterior data association probabilities are calculated to discriminate clutter measurements when updating trajectory probability density function. A new complete recursive track initiation, confirmation and deleting algorithm based on PF and Integrated Track Splitting (ITS) and named Integrated Particle Filter (IPF) is presented. Through the extended simulations showed the effectiveness of this approach in a five targets scenario.**

***Index Terms*—Target tracking, data association, particle filter, Integrated Track Splitting.**

I. Introduction

Each sensor measurements may either be a spurious (clutter) or a target measurement. The target existence and trajectory are not a priori known [1]. The tracks are initialized using measurements, thus both true tracks and false tracks simultaneously exist. The false track discrimination (FTD) is a procedure to terminate a majority of false tracks and confirm majority of true tracks [2],[3]. A track quality measure needs to be calculated for successful FTD. The multiple hypothesis tracker (MHT) [4][5] is one of the first widely used algorithm for target tracking in clutter. The measurement-oriented MHT, often known as the Reid algorithm [1], forms new tracks and measurement allocation hypotheses centered around global origin of measurements. The MHT uses statistical methods (track score) to discriminate between false and true tracks. The probability of target existence obtained by utilizing Markov chain propagation models and Bayes update is used as the track quality measure in Integrated Probabilistic Data Association (IPDA) of [6] and Integrated Track Splitting (ITS) [7],[8].

The application of the Sequential Monte Carlo estimation framework to real multi-target tracking problems is plagued by many difficulties. Among other things, realistic models for the target dynamics and measurement processes are often nonlinear and non-Gaussian, so that no closed-form analytic expression can be obtained for the tracking recurs.

When tracking a single object closed-form expressions are generally not available for nonlinear or non-Gaussian models, and approximate methods are required. The extended KF liberalizes models with weak nonlinearities around the current state estimate, so that the KF recursions can still be applied. However, the performance of the EKF degrades rapidly as the nonlinearities become more severe. To alleviate this problem the unscented KF (UKF) [9], [10] maintains the second-order statistics of the target distribution by recursively propagating a set of carefully selected sigma points [11]. This method requires no linearization, and generally yields more robust estimates.

When tracking with Particle Filter [12],[13] an analog to the predicted measurements covariance is not directly available and could only be constructed as an approximation to the current particle cloud. A common alternative is to use a form of soft gating based upon a Student'st likelihood, combine the same function and probabilistic data association approaches to develop a new method for tracking in clutter using a particle filter. This is done by deriving an expected likelihood from known measurements and clutter statistics.

In this paper, we propose the integrated particle filter (IPF) solution for the target tracking in clutter. Each track trajectory pdf is represented by a disjoint set of particles, and the probability of target existence is integrated into the track state, similar to [14], [15], [16]. The FTD may use the probability of target existence as the track quality measure. The standard IPF is a single-target tracker, and we also include multi target approach [17] for target tracking. They all share common recursion elements, being distinguished by the data association calculus. In addition to the recursive calculation of the probability of target existence and non-uniform clutter, we also include the state dependent probability of target detection, and maneuvering (multi-model) target trajectories [18].

Rest of the paper is organized as follows. The models and the particle filter background are presented in Section 2. The common IPF framework is detailed in Section 3, and the implementations of IPF is presented in Section 4. This approach is indicated by simulations in Section 5, followed by the concluding remarks in Section 6.

II. Problem Statements

The dynamic target trajectory state models at the time k are given by the:

$$x_k = Fx_{k-1} + \nu_k \tag{1}$$

where F is the propagation matrix, ν_k is a zero mean and

Branko Kovacević is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: kovacevic_b@etf.rs).

Zvonko Radosavljević and Dejan Ivkovic are with the Military Technical Institute of of Belgrade, Ratka Resanovica 1, 11000 Belgrade, Serbia (e-mail: zvonko.radosavljevic@gmail.com and divkovic555@gmail.com).

white Gaussian sequence with covariance *R*. At each scan the sensor returns a random number of random target and clutter measurements. The measurement of existing and detectable target is taken with a probability of detection P_D. At time *k*, one sensor delivers a set of measurements $z_k = \{z_{k,j}\}_{j=1}^{M_k}$ track out of which a set of measurements are selected for track update. Converted target measurement *y* is given by [19] :

$$y_k = Hx_k + w_k \tag{2}$$

where *H* is measurements matrix and the measurements noise w_k is zero mean and white Gaussian sequence. A measurements of target is present in each scan with a probability of detection P_D . Clutter measurements follow the Poisson distribution characterized at location by clutter measurements density $\rho_k(y)$ [19].

Particle filtering samples at the continuous posterior density function of interest into a set of weighted particles. If the weights are chosen appropriately, then these weighted set of particles represent the posterior density in a way that the posterior density function can be made arbitrarily close to the equivalent set of weighted particles. The target trajectory state *pdf* at scan *k* is defined by set of particles $\{x_k, w_k\}$, parameterized by set of *N* particles $\{w_l^i, x_k^i\}_{i=1}^N$ where should be satisfied $\sum_{i=1}^N w_l^i = 1$. Using sequential importance sampling [xx], particle filters can approximate the posterior density function, regardless of the time interval k of the trajectory model [20].

III. Integrated particle filter

The track state consists of the target existence event, and the trajectory state, and for each track we recursively calculate the probability of target existence, and the trajectory state probability density function (pdf). The trajectory state pdf are only defined conditioned on target existence. Depending on the calculated probability of target existence we may conclude that the target exists and confirm the track. Each confirmed track stays confirmed until termination. Alternatively, if the calculated probability of target existence dips below certain level we conclude that the target does not exist and terminate the track [21].

Key topics of new IPF algorithms are:

- new particles arise by re-sampling;
- heavy particles are multiply,
- weak particles are extinguished
- measurements are used to correct the weight of the particles and the probabilities of target existence.

At begin, lets define key parameters. The number of particles from $(k-1)^{th}$ scan, $N_{k-1} = N$ does not change from scan to scan. Lets represent particle $\{x_{k-1}^i, w_{k-1}^i\}$, $i = 1,..., N_{k-1}$ from $(k-1)^{th}$ scan , mean and weight. Number of measurements arriving from k^{th} scan are M_k, and $N_p = N$ is number of particles after re-sampling step. Probability of target detection, as the function of target trajectory state is $p_D(x_k) = P_D$. Also we have equation [22]:

$$\tilde{P}_D = \sum_i w_{k-1}^i p_D(x_k^i) = P_D \sum_i w_{k-1}^i = P_D \tag{3}$$

Proposed IPF algorithm is perform by the following steps:

- prediction step,
- measurements likelihood calculating
- update step and
- re-sampling step.

A. Prediction step:

At begin, we calculate probability of target existence, by the:

$$\psi_{k|k-1} = \Delta_{11} \cdot \psi_{k-1|k-1} \tag{4}$$

The mean of particle is given by the:

$$x_k^i = f(x_{k-1}^i, v_k^i) = Fx_{k-1}^i + v_k^i \tag{5}$$

where particle propagation noise is $v_k^i \approx N(0,Q)$ and measurements sets is given by $Z_k = \{z_k^1,..., z_k^{M_k}\}$

B. Measurements likelihoods

After KF prediction, we estimate measurements by the:

$$\hat{y}_k^i = Hx_k^i \tag{6}$$

In order to compute statistical distance:

$$d^2{}_{ij} = (z_{k,j} - \hat{y}_k^i)^T (R_k)^{-1} (z_{k,j} - \hat{y}_k^i), \; j = 1,..., M_k \tag{7}$$

Probability density function is given by the:

$$p_{k,j}^i = \frac{1}{\sqrt{\det(2\pi R_k)}} \exp[-0.5 \cdot d^2{}_{ij}] \tag{8}$$

where likelihoods of measurements is:

$$p_{k,j} = \sum_i w_{k-1}^i \cdot p_{k,j}^i \tag{9}$$

Now, we can calculate measurements likelihood ratio, by the equation:

$$\Lambda_k = 1 - P_D + P_D \sum_j \frac{p_{k,j}}{\rho_{k,j}} \tag{10}$$

Beta's coefficients we can update by the:

$$\beta_{k,j} = \frac{1}{\Lambda_k} \begin{cases} 1 - P_D, & j = 0 \\ P_D \dfrac{p_{k,j}}{\rho_{k,j}}, & j > 0 \end{cases} \tag{11}$$

C. Update step:

In update step, we first calculate weight of particles, in purpose of trajectory state update, by the [23]:

$$w_k^i = w_{k-1}^i \cdot (\beta_{k,0} + \sum_{j=1}^{M_k} \beta_{k,j} \frac{p_{k,j}^i}{p_{k,j}}) \tag{12}$$

At the end of update step, we calculate target existence probability of track, by the equation:

$$\psi_{k|k} = \frac{\Lambda_k \psi_{k|k-1}}{1-(1-\Lambda_k)\psi_{k|k-1}} \tag{13}$$

D. Resampling step:

Resampling step calculates mean and weight of particles, by the foloowing [24]:

$$\{x_k^i, w_k^i\} \Rightarrow \left\{ x_k^I, w_k^I = \frac{S_w}{N_p} = \frac{1}{N} \right\}, I = 1, 2, ..., N \tag{14}$$

where

$$S_w = \sum_{i=1}^{N_{k-1}} w_k^i = 1 \tag{15}$$

$$u_1 = U\left[0, \frac{1}{N}\right] \tag{16}$$

$$u_l = u_1^l + (l-1)\frac{1}{N}, i_c = i_{c-1} + w_k^i, i = 1, ..., N \tag{17}$$

where S_w is sum of weights, $U[.]$ means uniform distribution, u_l is interval of weights.

E. *Output Calculation*

Finally, we can calculate the output state estimate and covariance (for output purpose only):

$$\hat{x}_{k|k} = \sum_{l=1}^{Np} w_k^l x_k^l \tag{18}$$

$$P_k = \left(\sum_{l=1}^{Np} w_k^l \cdot x_k^l \cdot x_k^{l\,T} \right) - \hat{x}_k \cdot \hat{x}_k^{\,T} \tag{19}$$

IV. IMPLEMENTATION OF IPF

In this section, a brief instruction of IPF sofware implemenattion, we describe. Track initiation and termination is an part for establishing the records of the new targets and terminating the unwanted records of the inexistent targets when they leave the surveillance region. But in the heavy cluttered environment, there exists due to the unknown state of the target and the sequence of measurements which originate from the target. Here, we present a track management procedure.

Track initiation is composed of two parts:

- produce temporal tracks and
- confirm the temporal tracks.

Track termination is of two meanings:

- reject the temporal tracks;
- terminate the confirmed tracks when the detected targets leave the surveillance region.

A. Software implementation of IPF

One cycle of the recursive IPF algorithms software implementation consists of the following procedure:

for scan = 1 : number of scans
--Read Measurements
-Target Tracking with IPF
-Initializing of Measurements Selection
-Measurements selection (measurement likelihood for all particles)
-Taking into account clutter density
-Update Tracks of IPF
-Single Target Track Data Association
-Update Weights
-Resampling
-Estimate IPF
-Tracks Initializing
-Update Old Samples
-Update Status
-Update Age
-Eliminate Wide
-Merge Close Tracks
-Eliminate Tracks -Out of Bound
-Update Tracks (Confirmation and Deleting)
-Prediction of IPF
-Determine Target Track
-Target Statistics of Scans (True, False, Confirmed,…)
-Reduce Tracks
End

V. SIMULATIONS

For the purpose of research, a simulation scenario with five targets motion scenario (Fig.1). Targets are initially positioned at the edges of a circle with the center at (500,500) and a radius of 450. Each target moves with a uniform speed towards the center of the circle, which they should reach in 20 scans, after which they carry on with uniform motion for further 20 scans. A random (noise) component is added to the speed vector of each target, with covariance (2*R/400).

A random component is added to the speed vector of each target, thus at scan 20 the variance of the distance between each target and the centre of the circle will be double the sensor measurement error noise covariance matrix. In the two targets scenario, the targets initial separation is 20^o, instead of fifteen targets scenario with the targets initial separation 10^o. The following definitions of true and false tracks are used. Each initiated track is false with respect to all existing targets. A false track becomes a true track with respect to a target when the state estimate is sufficiently close to the true target state.

Each simulation experiment consists of a number of simulation runs. In each simulation run, targets will repeat their trajectories. The measurements are generated independently. Each algorithm uses the same set of measurements. False tracks may be initiated using target measurements, either in a conjunction with a clutter measurement, or by using measurements from different targets in different scans.

Thus, the average number of initialized false tracks per scan will depend on the number of targets present. The average number of initialized false tracks per scan was 8, and 120 for the two and fifteen targets experiments, respectively. A confirmed false track in one scan is 300 and 200 for the two and fifteen targets cases, respectively. The performance measures used to compare the algorithms confirmed true tracks, root mean square error positions and target retention statistics. Results are presented by a number of confirmed true tracks and Root Mean Square Error Position.

The target retention statistics was obtained by noting the identity of the confirmed true track following each of the targets at scan 14. These identities are checked again at scan 38, and the following statistics is accumulated for each experiment:

nCases: total number of cases of a target being followed by a confirmed track at scan 14;

nOK: percentage of tracks still following the original target at scan 38;

nSwitched: percentage of tracks that end up following a different target at scan 38;

nLost: percentage of tracks not following any target at scan 38,

nMerged: percentage of tracks lost due to merging between tracks counted in nCases between scans 14 and 38

For the target retention statistics, each algorithm identifies the confirmed true track for a specific interval that includes intersection of trajectories. The targets intersect at scan 24 and many joint events occur around that time. In the experiment, the identities of the confirmed true tracks are obtained at scan 17 for performance comparison.

Parameters were used: probability of target detection is pD=0.8, number of Monte Carlo runs is 100, duration of one recursion 40, measurements noise matrix -R=[25 0; 0 25], maximum of target speed -25 [m/s], variance of acceleration q =0.75, number of particles -1000, maximum number of components -40, starting cross statistics in 14 scan, ending cross ststistics scan 38.

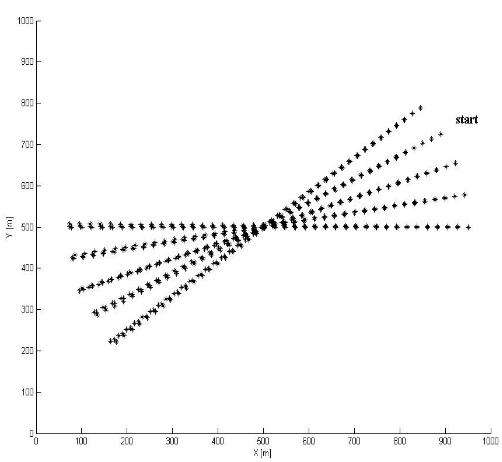


Fig. 2. Simulation scenario (Five targets)

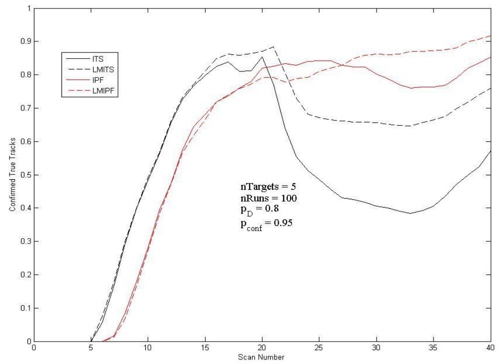


Fig. 2. confirmed true tracks diagram over time (five target)

The sampling period of radar sensor is T=1s. Duration of the scenario is 40 scans. The system input is modeled as follows: vector state $\mathbf{x}(k) = [x\ \dot{x}\ y\ \dot{y}]^T$ where the Cartesian coordinates of the target position are, and are the appropriate velocities. Transition matrix and process noise matrix are given by:

$$F = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{20}$$

$$Q_k = q\begin{bmatrix} T^3/3 & T^2/2 & 0 & 0 \\ T^2/2 & T & 0 & 0 \\ 0 & 0 & T^3/3 & T^2/2 \\ 0 & 0 & T^2/2 & T \end{bmatrix} \tag{21}$$

respectively. Measurements matrix and measurements noise matrix is given by:

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix},\ R_k = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \tag{22}$$

respectively.

All simulations were done in a software package MATLAB, with CPU Intel Core i7, 2.93 GHz. Results of simulation are governed by the number of confirmed true tracks (Fig, 2) and target retention table. We compare standard ITS and proposed IPF algorithms.

Target retention table

	ITS	**IPF**
nCases[n]	**91**	**80**
nOK[%]	31.86	42.5
nSwit[%]	15.38	18.75
nLost[%]	52.76	38.75
merged	26	14
CPU [s]	1.65	1.81

The results confirm the justification of the proposed IPF approach compared to standard ITS algorithm. IPF has a smaller percentage of losses and switched targets and higher percentage of full tracking targets with approximately the same CPU consumption.

VI. Conclusion

The multiple target tracking algorithm, known IPF, is proposed and was tested in a special scenarios with five crossing targets. It uses the well-known features of ITS algorithms that account the probability of target existence of objective forms, trace and ease of use offered by the Particle Filter. A Simulation results with two-dimensional scenario showed that the proposed algorithm ends up with good performance and small computational load. Proposed algorithm, which has been presented for tracking multi, have the ability to estimate the number of targets. Tracking the trajectories of the target over time, operate with missed detections and give the trajectories of the targets.

References

[1] D. Reid, 'An algorithm for tracking multiple targets', IEEE Trans. AC, vol. AC-24, no. 6, pp. 843–854, 1979.

[2] Blackman, S.:'Multiple-target tracking with radar applications' (Artech House, 1986)

[3] Kurien, T.:'Issues in the design of practical multitarget tracking algorithms', IEEE Trans. Autom. Control, 1979, 24, (6), pp. 843–854

[4] Challa, S., Evans, R., Morelande, M., Mušicki, D.: 'Fundamentals of object tracking' (Cambridge University Press, 2011)

[5] Musǐcki, D., Evans, R., Stankovic´, S.: 'Integrated probabilistic data association (IPDA)', IEEE Trans. Autom. Control, 1994, 39, (6), pp. 1237–1241

[6] Mušicki, D., La Scala, B.: 'Multi-target tracking in clutter without measurement assignment', IEEE Trans. Aerosp. Electron. Syst., 2008, 44, (3), pp. 877–896.

[7] Bar-Shalom, Y., Tse, E.: 'Tracking in a cluttered environment with probabilistic data association', Automatica, 1975, 11, pp. 451–460

[8] Mušicki, D., Evans, R.: 'Multiscan multitarget tracking in clutter with integrated track splitting filter', IEEE Trans. Aerosp. Electron. Syst., 2009, 45, (4), pp. 1432–1447

[9] Z. Radosavljević, T.L. Song, B.Kovacevic, Linear Multi-Target IPF Algorithm for Automatic Tracking, Scientific Technical Review, 2016,Vol.66,No.1,pp.3-10, Belgrade 2016.

[10] Bar-Shalom, Y., Li, X.R., Kirubarajan, T.: 'Estimation with application to tracking and navigation' (John Wiley and Sons, 2001)

[11] Julier, S.J., Uhlmann, J.K.: 'Unscented filtering and nonlinear estimation', Proc. IEEE, 2004, 92, (3), pp. 401–422

[12] Alspach, D.L., Sorenson, H.W.: 'Nonlinear Bayesian estimation using Gaussian sum approximation', IEEE Trans. Autom. Control, 1972, 17, (4), pp. 439–448

[13] Ristic, B., Arulampalam, S., Gordon, N.: 'Beyond the Kalman filter' (Artech House, 2004)

[14] Hue, C., Le Cadre, J., Pérez, P.: 'Sequential Monte Carlo methods for multiple target tracking and data fusion', IEEE Trans. Signal Process., 2002, 50, (2), pp. 309–325

[15] Marrs, A., Maskell, S., Bar-Shalom, Y.: 'Expected likelihood for tracking in clutter with particle filters'. SPIE: Signal and Data Processing of Small Targets, April 2002, vol. 4728, pp. 230–239

[16] Vermaak, J., Godsill, S.J., Pérez, P.: 'Monte Carlo filtering for multi-target tracking and data association', IEEE Trans. Aerosp. Electron. Syst., 2005, 41, (1), pp. 309–332

[17] Song, T.L., Mušicki, D.: 'Adaptive clutter measurement density estimation for improved target tracking', IEEE Trans. Aerosp. Electron. Syst., 2011, 47, (2), pp. 1457–1466

[18] Gordon, N.J., Salmond, D.J., Smith, A.F.M.: 'Novel approach to nonlinear/ non-Gaussian Bayesian state estimation'. IEE Proc., Radar, Sonar and Navigation, 1993, vol. 140, no. 2, pp. 107–113

[19] Song, T.L., Mušicki, D., Kim, D.S.: 'Target tracking with target state dependent detection', IEEE Trans. Signal Process., 2011, 59, (3), pp. 1063–1074.

[20] Song T.L., Mušicki D., Kim D.S., and Radosavljević Z.: 'Gaussian mixtures in multi-target tracking: a look at Gaussian Mixture Probability Hypothesis Density and Integrated Track Splitting', IET Proceedings: Radar, Sonar and Navigation, vol. 6, 2012, no. 5, pp. 359-364..

[21] Radosavljević Z., Mušicki D.: Limits of target tracking in heavy clutter, ASIA-Pacific International Conference of Synthetic Aperture Radar APSAR 2011, Seoul, Korea, 2011.

[22] Z .Radosavljevic, D. Ivkovic, 'An Approach of Track Management System in Software Defined Radar', Scientific Technical Review, 2018, Vol.68,No.1, pp.73-80, Belgrade 2018, UDK 681:5.017:623:746.3

[23] Z. Radosavljević, ,IPDA Filters in the Sense of Gaussian Mixture PHD Algorithm, Scientific Technical Review, 2016,Vol.66,No.3, pp.34-40, Belgrade 2016, UDK : 681:5.017:623:746.3, COSATI:17-09, 12-01.

[24] Radosavljević, Z., Musǐcki, D., Kovačević, B., Kim, W.C., Song, T.L.: 'Integrated particle filter for target tracking'. 13th Int. Conf. on Electronics, Information and Communication, ICEIC 2014, Kota Kinabalu, Malaysia, January 2014.

Hough transform in visual product quality control

Aleksandra Marjanović, Sanja Vujnović and Željko Đurović

***Abstract*— Product quality inspection is one of the indispensable steps in the production process, and there are more and more factories that are trying to automate that procedure by using computer vision algorithms. Additional efforts are made to keep these algorithms simple and fast when time is of the essence. This paper relies on Hough transform as a standard tool in image processing and discusses its possibilities in a time-constrained scenario. Being that the considered product is ball-shaped, the extension of Hough transform for circle detection is used to detect product in appropriate cells on the conveyor belt. The problem setup may seem easy, but unpredictable parameters of the industrial surroundings make it challenging. The detection algorithm is tested on a real-life image database collected at one chemical factory in Serbia.**

***Index Terms*—Visual product quality control, Defect detection, Hough transform, Circle detection.**

I. Introduction

With the development of modern technology, the efficiency and reliability of industrial plants is increasing, whether it is in terms of improving the hardware of existing systems or in terms of applying intelligent control laws that can monitor and regulate a large number of signals simultaneously. The automation of the production process is especially important in places that do not represent an ideal working environment for humans, such as plants in the electric power and chemical industries.

In the last decades advanced algorithms have found their place not only for increasing the quality of the production process, but also because of high expectations from customers, fierce competition, and stricter requirements of regulatory bodies and in quality control of the final product [1]. One of the basic forms of quality control is visual inspection of products and in many factories it is carried out in an old-fashioned way, by a human inspector. This, however, can be a very demanding job that requires a person to be in constant focus during a shift of about 8 hours, looking at the same product thousands of times. Research shows that in this process the error rate is high and goes over 25%. It is only natural that automated solutions for visual inspection of product quality, which rely on the methods of computer vision and artificial intelligence, are becoming more and more common [2], [3]. Apart from lowering the errors in the quality control process, there is another great advantage of implementing these kinds of solutions by the manufacturer. Mainly, the redirection of people from a repetitive non-creative job, which in some industries can be unpleasant for people's physical health, to jobs that are of greater importance, and which could not be done without people [4]. According to Markets Insider estimates, the machine vision sector will have earnings of about $ 12.29 billion from the start of their use until 2023, with an annual growth rate of about 7.61%.

The requirements placed on such a quality control system are very often contradictory. On the one hand, they must be fast enough for the production system to run smoothly, and on the other hand, more advanced algorithms that would increase the accuracy of such systems are also numerically more complex and require more time to execute. This paper discusses a segment of the visual inspection system in a modern chemical production plant that operates with a large capacity of several hundred products per minute. This leaves room for about a hundred of milliseconds for product processing, which should include image recording, image processing and decision making, as well as taking appropriate action. The job of engineers, therefore, is to design the simplest possible decision system with the highest possible success rate. Of course, it should be emphasized that the mechanical part of the system should be fast and reliable, but at the same time simple and cheap enough to be cost-effective for serial production.

The problem tackled in this paper refers to quality control in the production process of scented sanitary balls. Prior to the automation of this process, the inspection was performed by a person by monitoring for several hours the batches of balls that are at the output of the shaping subsystem. Apart from the fact that this kind of work is tiring for the sense of sight, it is also tiring for the sense of smell due to strong vapors in the plant. Due to all these facts, the company's management came up with the idea of automating this segment of production. This paper describes the first step in the product quality assessment procedure, which refers to the detection of products within the intended slot. In real industrial conditions this can be quite challenging, as will be explained later. The algorithm must be simple and fast enough at the same time, so that in combination with the algorithm for checking the regularity of the ball and communication between individual components it is suitable for real time application.

Marjanović is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: amarjanovic@etf.bg.ac.rs).

Sanja Vujnović is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: svujnovic@etf.bg.ac.rs).

Željko Đurović is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: zdjurovic@etf.bg.ac.rs).

Following Introduction, Section II provides a review of Hough transform and its application in line and circle detection. Section III describes in detail the setup of product visual inspection process, focusing on the challenges during product detection phase and the possibilities of overcoming them using Hough transform. Finally, Section IV shows main results concerning the influence of image down sampling on detection accuracy, as a tool for reduction of computational efforts.

II. HOUGH TRANSFORM

Hough transform in its original form is a tool that enables fast and efficient detection of straight lines in digital images [5]. Namely, if we want to find among the n points in the plane those that belong to one line, the number of operations that should be done by the brute-force method is proportional to n^3, which is unusable in the context of real time execution. On the other hand, Hough transform starts from the idea that one point (x_i, y_i) located on the line

$$y = ax + b \tag{1}$$

maps to the line in (a, b) space

$$b = -x_i a + y_i. \tag{2}$$

Similarly, any other point on the line (1) will map into the line in (a, b) space, which intersects the line (2) in the point determined by parameters a and b in Eq. (1). In other words, the problem of searching for line parameters is reduced to searching the space (a, b) with the aim of finding the point where the largest number of lines intersect [6]. A limiting circumstance in this consideration is the fact that the vertical line has an infinite slope. This problem is overcome starting from the idea that every point on the line in (x, y) space can be represented by a sinusoid in the parametric space (ρ, θ)

$$x \cos\theta + y \sin\theta = \rho, \tag{3}$$

which is shown in Fig. 1.

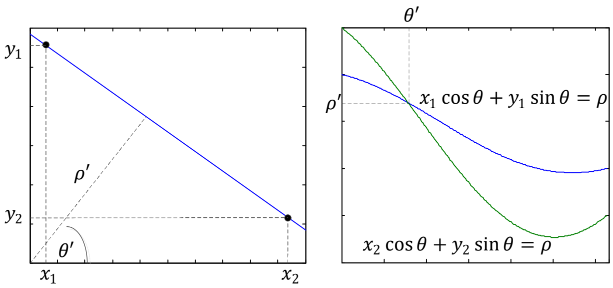


Fig. 1. The concept of Hough transform.

Similarly, the cross section of these sinusoids contains the correct parameters ρ and θ. In this way, the previous search is reduced to a search by angle $\theta \in [-\pi/2, \pi/2]$ and $\rho \in (-\infty, \infty)$. The implementation of this algorithm first involves the quantization of space (ρ, θ) into so-called accumulator cells. Then, each point (x_k, y_k) is observed and the corresponding quantized value ρ is calculated for each quantized value of the parameter θ and the final search for the cell in which the most sinusoids were found. Over time, the extensions of this procedure to the problem of finding second-order curves and arbitrary shapes have been considered [7], [8]. Thus, for the detection of a circular contour, three parameters should be determined, the coordinates of the center of the circle and the radius:

$$(x - x_0)^2 + (y - y_0)^2 = R^2. \tag{4}$$

Similarly as before, a different representation of the circle over the radius and polar angle can be considered:

$$\begin{aligned} x &= x_0 + R\cos\theta \\ y &= y_0 + R\sin\theta, \end{aligned} \tag{5}$$

where $\theta \in [0, 2\pi)$. In this case, the accumulator is three-dimensional, and the procedure for incrementing the cell values is based on the previously described procedure. Namely, each pair of points is a potential center (x_0, y_0), and the radius range is the potential radius. Bearing in mind that the search is now done in three-dimensional space, computational time of single curve detection is high. Over time, improvements to this algorithm have been proposed, such as the use of genetic algorithms [9] and harmony search [10], which shorten computational time and allow sufficiently precise circuit parameters to be found. Also, there are extensions of this algorithm to the detection of arbitrary curved lines.

III. PRODUCT DETECTION ALGORITHM

To better understand the idea of applying Hough transform in the process of visual quality inspection, let us describe the setting of the problem in a little more detail. The balls arrive one by one for individual inspection, sorted on the appropriate conveyor belt, separated by barriers (Fig. 2). The part of the system related to the visualization of the product consists of three cameras separated in space, two on the side and one from above, which have the task to look from different angles at the cell in which the product should be located. After that, in case there is a product in the cell, a quality control algorithm is applied, which should assess whether the ball is defective or not. If there is no defect, the ball should be passed to the next step of the production process. If, however, the ball is defected the pneumatic blower should be signaled to throw the ball off the production line, after which it is sent for recycling, i.e., the beginning of the mixture making step. Processing images from two side cameras can be done in the same way, because the image obtained from them is similar. Namely, it shows a ball leaning on the barrier that is immediately behind the ball when the direction of movement of the conveyor belt is observed, and in one case the ball is leaning on the left and in the other on the right barrier.

Additionally, behind the ball (as seen from the camera), there is a partition that provides an appropriate, contrasting background color depending on the color of the ball. On the other hand, the third camera is placed above the ball, and the picture from it shows the ball leaning on the barrier, but also a part of the base, that is. the conveyor belt on which the ball lies. Precisely because of this difference, processing images from the third camera is somewhat more complicated than the first two. Namely, after a few minutes, and then a few hours of using the system, the belt becomes dirty, sometimes with the color of the balls that previously passed over it, which is shown in Fig. 3 (one and the same belt is often used to process different colors of balls, depending on the requirements). In order for the belt to be cleaned, the machine must be stopped for a few minutes. However, if the cleaning is not carried out in a timely and adequate manner, this can complicate the algorithm for processing images from the third camera.

The step of checking whether there is a ball between the two barriers is the first step in the algorithm, and it should be the fastest one, so that there is enough time to perform a more time-consuming step, which refers to checking the quality of the ball. By analyzing the image from the first two cameras, this is very easy to check, because the background behind the ball is not dirty. However, if the same algorithm is applied to images from the third camera, many false alarms appear, which further leads to a misconception about the number of defective balls, because empty slots are classified as irregular balls in the next step. An additional problem in the analysis of the image from the third camera is the fact that during operation the voltage on the light is variable, so the brightness of the ball changes with it, and it is difficult to automatically adjust the camera exposure, because it is almost impossible to find a reference point that will not be influenced by the color of the ball. These and similar problems are precisely the consequence of working in industrial conditions, which can be very unpredictable.

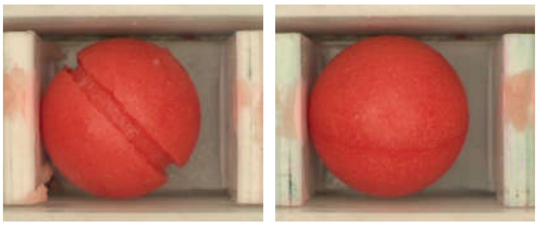


Fig. 2. Examples of full cells.

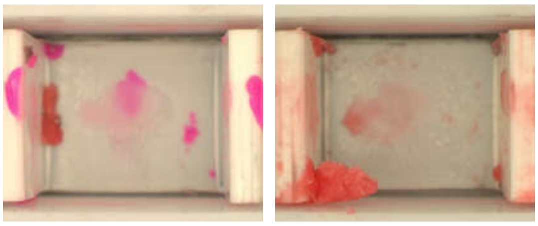


Fig. 3. Examples of empty cells.

This is exactly the problem we will try to solve by applying Hough transform. Because the picture of an empty cell often consists of stains at the bottom of conveyor belt, as well as smaller masses which should not be counted as the ball and will certainly be eliminated from further production process due to smaller dimensions, Hough transform is proposed to find circular contours in an image whose radius is in a predefined range, and in accordance with what the expected radius of the ball is. The algorithm consists of several steps:

- In the first step, depending on the color, it is necessary to find the appropriate color space, i.e. the appropriate representation, in which it will be easier to notice the difference between the ball and the background. In some cases, it will be some of the RGB components, other times HSV space will be more useful, etc.
- In the next step, using the Canny edge detector, the previously selected gray image is translated into a black and white image, by carefully selecting the thresholds of the said detector. The thresholds are chosen so that the edges of the ball are detected as clearly as possible, although in real conditions this implies the detection of background noise.
- Finally, the Hough transformation is applied to check whether among the detected points in the black and white image there are those that belong to the same circular contour of the appropriate size, which indicate the presence of a ball in the image. Due to the nature of the algorithm, it will certainly find a circle, and based on the parameters of the circle and the number of detected points, a decision is made whether the ball exists in the image or not.

Additionally, as mentioned earlier, this procedure involves searching in 3D space, which requires computing resources. Bearing in mind that the ball has some expected size, i.e. the radius, as well as the limited space in which the center of the sphere can appear, the first step in enhancing the speed of the algorithm refers to the appropriate narrowing of the search space for all three parameters. The second step refers to the examination of the extent to which it is possible to work with a resampled image of lower resolution, i.e. the extent to which the decimation procedure affects the finding of the circular contour and its parameters. The results of this consideration are presented in the next chapter.

IV. RESULTS

Testing of the proposed algorithm was done on a database consisting of 3000 images with empty cells and 3000 images with the product (2800 regular products and 200 defects). All images have the same 600x500 resolution. Let us first observe the performance of circular contour detection in different images. Fig. 4 and Fig. 5 clearly show that the extended detection algorithm gives a good result, in terms of successful contour detection. The images on the left show the original images of the balls, and on the right is the result of the application of Canny edge detector, together with a drawn contour whose parameters correspond to a circle that fits the largest number of points. Fig. 6 shows the results in the case of images without products. What can be noticed at first glance is that the algorithm again managed to find a circular contour, although it is not obvious in the original image. But that is still the expected result, because the algorithm finds the circle on which the largest number of points from the BW image is located. The essential difference between these two cases is in the number of points which the detected circle fits.

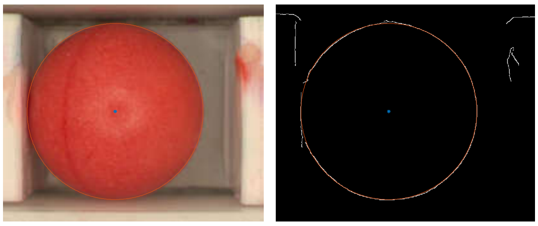


Fig. 4. Examples of circle detection on a regular product.

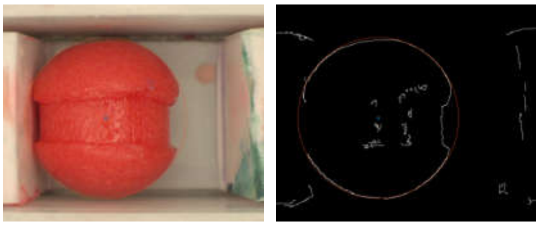


Fig. 5. Examples of circle detection on a defect.

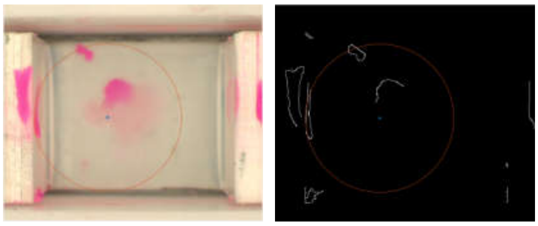


Fig. 6. Examples of circle detection on an empty cell.

Table 1 shows first- and second-order statistics for the number of detected points per circle N, as well as the execution time of the algorithm T on the Intel(R) Core(TM) i7-9700KF CPU at 3.60 GHz configuration It can be observed that the number of detected points is significantly higher in the case when the ball exists than when the cell is empty. A more detailed analysis showed that there is a clear gap between these two cases, which means that the detection problem can be successfully solved by setting an appropriate threshold. In addition, another interesting fact is that in the case off a defect, the number of detected points is slightly smaller than in the case of regular products. This result is justified bearing in mind that the defect of the product is often reflected in its irregular contour, which deviates from the circular contour in a few segments. However, for most products, this deviation is not significantly pronounced, which is why these two cases are not linearly separable considering this parameter space. Execution time analysis in different cases also gives the expected results. Namely, in the case of a defective ball, the output from the Canny edge detector is slightly higher than in the case of a regular one. Therefore, the number of points that need to be processed during the formation of the accumulator matrix is higher. On the other hand, one could expect that in the case of empty cells, the number of points returned by the Canny edge detector is not large. However, depending on the lighting and the degree of conveyor belt contamination, the number of points may be approximately the same as in the case of regular balls. Therefore, the execution time of the algorithm in that case is somewhat less, but not significant. Even though 30ms does not sound like a long time, it is not affordable when there is around 150ms available for the overall product inspection process.

TABLE I
AVERAGE DETECTION PARAMETERS PER DIFFERENT TYPES OF IMAGES

	N_{mean}	N_{std}	T_{mean}	T_{std}
Regular product	312.6	76.3	29.5ms	2.1ms
Defect	265.9	65.9	32.1ms	3.3ms
Empty cell	41.5	13.17	26.8ms	4.2ms

The last result led us to the idea of decimation, i.e. downsampling. Several cases depending on the degree of decimation were considered. Table 2 presents the results that show the degradation of detection parameters depending on the degree of decimation. In addition to the two parameters discussed earlier, this analysis also observed the extent to which the position of the center and the radius of the detected circle change with decimation in comparison to the case of no decimation. What can be noticed from the last two columns is that the relative changes in the position of the center and the radius are very small. What is perhaps a slightly more indicative information is that the maximum absolute deviation in the center position in the case of an empty cell is between 80 and 90 pixels, and in the case of a full cell between 18 and 22 pixels depending on the parameter k. Similarly, during decimation, there is a maximum absolute change in the radius of 12-18 pixels for full cells, and 26-28 pixels for empty cells. In other words, the difference obviously exists, but it is negligible when compared to the radius. Another important parameter is the mean execution time of the algorithm. The results show that this parameter is smaller for higher k, and that it decreases by about 30% for higher decimation.

TABLE II
AVERAGE DETECTION PARAMETERS FOR DIFFERENT DEGREE OF DECIMATION

		N_{mean}	T_{mean}	$\Delta x/R$	$\Delta R/R$
$k=2$	Full	156.9	24.1ms	4.4e-3	0.7e-3
	Empty	22.9	21.8ms	25.3e-3	9.0e-3
$k=3$	Full	103.9	21.4ms	3.3e-3	1.5e-3
	Empty	14.6	20.2ms	26.4e-4	11.5e-3
$k=4$	Full	79.7	20.2ms	6.5e-3	2.2e-3
	Empty	11.9	19.3ms	36.4e-3	16.3e-3
$k=5$	Full	62.9	19.7ms	5.1e-3	4.6e-3
	Empty	9.3	18.9ms	35.9e-3	19.3e-3

Based on previous considerations, the introduction of decimation is justified. Let us analyze how this procedure affects the number of points N as a crucial parameter for decision making. The mean value of the number of points on the detected circle decreases with a higher degree of decimation, which is expected. Again, when looking at the mean value in these two cases, it would seem that there is a large enough gap between the classes, however it turns out that they are closer to each other with higher decimation, primarily due to irregular products. Therefore, although the previous arguments are in favor of a higher degree of decimation, one must still be cautious and a compromise must be made. In this sense, it is shown that for $k = 4$ it is still possible to set a threshold on the number of detected points for classification purposes, and yet significantly speed up the execution of the algorithm.

V. Conclusion

The paper discusses the possibilities of applying Hough transform in detection of a ball-shaped product in the appropriate cell on the conveyor belt during the product quality control, based on the digital image. Firstly, it was shown that images without a product can be classified quite easily from the case when the cell is full by setting a decision threshold. The possibilities of the BW image decimation procedure for the purpose of accelerating the detection algorithm, are further discussed. This step was evaluated by the execution time and individual parameters of object detection in relation to the case when no decimation exists. It has been shown that the execution time is significantly reduced to about 20 ms, which is about 30% shorter than in the original case. In terms of the parameters of the detected circle, the algorithm is not too sensitive, while the unfavorable influence of the degree of decimation is reflected in the reduction of the gap between the two classes in terms of the number of detected points on the circle. As a compromise solution, the decimation parameter $k = 4$ is proposed, for which the proposed algorithm does not make any mistakes during the decision-making process on the available database.

Acknowledgment

This research was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

References

[1] T. Johnson, S. Fletcher, W. Baker, R. Charles, "How and why we need to capture tacit knowledge in manufacturing: Case studies of visual inspection," *Applied ergonomics*, vol. 74, pp. 1-9, 2019.

[2] A. Djavadifar, J. Graham-Knight, M. Körber, P. Lasserre, H. Najjaran, "Automated visual detection of geometrical defects in composite manufacturing processes using deep convolutional neural networks," *Journal of Intelligent Manufacturing*, pp. 1-19, 2021.

[3] T. Wang, Y. Chen, M. Qiao, M, H. Snoussi, "A fast and robust convolutional neural network-based defect detection model in product quality control," *The International Journal of Advanced Manufacturing Technology*, vol. 94, no.9, pp. 3465-3471, 2018.

[4] D, Brückner, "Innovating quality control mechanisms in aseptic drug manufacturing by means of isothermal microcalorimetry and tunable diode laser absorption spectroscopy," *Diss. University_of_Basel*, 2017.

[5] P. Hough, "Method and Means for Recognizing Complex Patterns," U.S. Patent *3 069 654*, 1962.

[6] R. Duda, P. Hart, "Use of the Hough Transformation to Detect Lines and Curves in Pictures," *Communication of the ACM*, vol. 15, no 1, pp. 11–15, 1972.

[7] D. Ballard, "Generalizing the Hough transform to detect arbitrary shapes," *Pattern Recognition*, vol. 13, no. 2, pp.111-122, 1981.

[8] H. Yuen, J. Princen, J. Illingworth, and J. Kittler, "Comparative study of Hough Transform methods for circle finding," *Image and Vision Computing*, vol. 8, no. 1, pp. 71–77, 1990.

[9] V. Ayala-Ramirez, C. Garcia-Capulin, A. Perez-Garcia, R. Sanchez-Yanez, "Circle detection on images using genetic algorithms," *Pattern Recognition Letters*, vol. 27, no. 6, pp. 652–657, 2006.

[10] E. Cuevas, N. Ortega-Sanchez, D. Zaldivar, M. Perez-Cisneros, "Circle detection by harmony search optimization," *Journal of Intelligent and Robotic Systems: Theory and Applications*, vol. 66, no. 3, pp. 359–376, 2012.

Some new results on stability of incommensurate fractional systems and their $\mathscr{L}_p$-norms

Rachid MALTI
Université de Bordeaux – IMS UMR 5218 CNRS, France
Email: firstname.lastname@ims-bordeaux.fr

***Abstract*—Stability of fractional systems is yet an open problem especially when dealing with incommensurate differentiation orders. The objective of the invited talk is twofold. First of all, a new method [1] for determining stability regions, in the parametric space, of fractional incommensurate systems is presented. It is based on interval arithmetics and allows, beyond the stability property, to specify the regions in the parametric space that have the same number of unstable poles. Hence, all transfer functions which parameters belong to the same stability region have the same stability property.**

Contrary to rational systems, the stability property of fractional systems does not guarantee the existence (or the boundedness) of the $\mathscr{L}_p$-norms, $1 \le p \le \infty$, of their impulse response. Hence, the second objective of the talk is to examine the existence conditions of these $\mathscr{L}_p$-norms [2]. The established results are used to choose a performance index for evaluating stable feedback control system performances.

I. Introduction

Fractional systems has been attracting a lot of interest during the last two decades in different fields of engineering and science, since the seminal work by Oldham and Spanier [3], [4] for modeling diffusive phenomena.

Fractional systems can be described in transfer function form:

$$F(s) = \frac{\sum_{i=0}^{M} b_i s^{\beta_i}}{1 + \sum_{j=1}^{N} a_j s^{\alpha_j}} \tag{1}$$

where the exponents of s can be ordered $0 < \alpha_1 < \alpha_2 < \ldots < \alpha_N$, $0 \le \beta_0 < \beta_1 \ldots < \beta_M$ and $(a_i, b_j) \in \mathbb{R}^2, \forall i = 0, 1, \ldots, M, \forall j = 1, 2, \ldots, N$. The multivalued function $s \mapsto s^\nu$ becomes holomorphic in the complement of its branch cut line of the complex plane, chosen to be along the negative real axis, $\mathbb{R}_{\le 0}$, including the branching points 0 and ∞. Hence, $F(s)$ is a meromorphic function in the complement of $\mathbb{R}_{\le 0}$ of the complex plane: $\mathbb{C} \setminus \mathbb{R}_{\le 0}$.

The $\mathscr{L}_p$-norm of $f(t)$, the impulse response of $F(s)$, is defined as[1]

$$\begin{cases} \|f\|_p = \sqrt[p]{\int_0^{+\infty} |f(t)|^p \, dt} & \text{for } 1 \le p < \infty \\ \|f\|_\infty = \sup_{t \in \mathbb{R}_{\ge 0}} |f(t)| \end{cases} \tag{2}$$

[1]Note that $f(t)$ is a continuous-time function and hence the supremum is used instead of the essential supremum in the definition of $\|f\|_\infty$.

A function $f(t)$ is said to belong to the Lebesgue space $\mathscr{L}_p(\mathbb{R}_{\ge 0})$ with $p \in [1, \infty]$, or $\mathscr{L}_p$ in short, if its p-norm is finite: $\|f\|_p < \infty$.

The system $F(s)$ is $\mathscr{L}_p$-stable, $1 \le p \le \infty$, if and only if:

$$\sup_{u \in \mathscr{L}_p, u \ne 0} \frac{\|f \star u\|_p}{\|u\|_p} < \infty \tag{3}$$

where $\star$ stands for the convolution product and $u(t)$ the system input. Condition (3) is satisfied, $\forall p \in [1, \infty]$, when

$$f \in \mathscr{L}_1(\mathbb{R}_{\ge 0}) \tag{4}$$

In such a case:

$$\|f \star u\|_p < \|f\|_1 \|u\|_p \tag{5}$$

The bounded-input-bounded-output (BIBO) stability is defined as the $\mathscr{L}_\infty$-stability.

Due to its simplicity, the most used criterion for testing stability of fractional systems is Matignon's theorem [5, theorem 2.21]. It allows deciding whether a system is stable by locating its s^ν-poles. It generalizes the classical Routh-Hurwitz criterion for rational systems. It is extended to take into account variations of $\nu \in (0, \infty)$ in [6].

Theorem 1. *A commensurate transfer function, of order ν, $F(s) = \frac{T(s^\nu)}{R(s^\nu)}$, where T and R are coprime polynomials, is BIBO-stable if and only if*

$$0 < \nu < 2 \tag{6}$$

and for every $s \in \mathbb{C}$ such that $R(s) = 0$

$$|\arg(s)| > \nu \frac{\pi}{2} \tag{7}$$

However, Matignon's theorem applies only for stability checking of commensurate fractional systems. When the system is incommensurate, some other criteria, mainly based on Cauchy's principal theorem [7] or its derivatives such as the Nyquist theorem [8] are used. However, these methods are quite difficult to implement in practice.

Since (4) is only a necessary condition, the system might be $\mathscr{L}_p$-stable and yet have an impulse response, f, with an infinite $\mathscr{L}_p$-norm $\forall p \in [1, \infty]$. This feature is not common in the classical rational systems: when a rational system is $\mathscr{L}_p$-stable then its $\mathscr{L}_p$-norms are finite.

II. Stability of uncommensurate fractional systems

This section of the paper contains results of a joint work with Milan R. Rapai'c and Vukan Turkulov, which is currently submitted at the ICFDA'2021 conference [1].

Bonnet and Partington [9] proved that $F(s)$ is BIBO-stable if and only if it is analytic in the right-half complex plane $\{\mathscr{Re}(s) \geq 0\}$. In such a case

$$F(s) \in \mathscr{H}_\infty(\mathbb{C}_+) \tag{8}$$

where $\mathscr{H}_\infty(\mathbb{C}_+)$ is the Hardy space of analytic functions on the open right-half plane $\mathbb{C}_+$.

On the other hand, it can be proven from Rouché's theorem that all system poles vary continuously with its parameters anywhere in the complex-plane except the branch-cut, $\mathbb{R}_{\leq 0}$. Moreover, when transfer function differentiation orders vary, new poles can appear or vanish only on the branch-cut. Hence, the basic idea is to consider that if a fractional system is stable for a given parametric point then it remains stable unless its poles cross the imaginary axis. Further, assuming that there is no possible simplification between poles and zeros of $F(s)$ in (1), the stability of $F(s)$ depends only on the position of the zeros of the characteristic function:

$$\bar{f}(s, \boldsymbol{\alpha}) = 1 + \sum_{j=1}^{N} a_j s^{\alpha_j} \tag{9}$$

A. Problem formulation

According to the previous remarks, the following problems can be formulated in the parametric space:

(P1) Finding stability and instability regions. In this case, the objective is to check whether, for positive $\boldsymbol{\alpha} \in \mathbb{R}^n_{\geq 0}$, $\bar{f}(\rho e^{j\theta}, \boldsymbol{\alpha})$ has zeros in the right half complex plane including the imaginary axis. However, due to the symmetry of complex conjugate zeros, the searching domain can be restreint to the first quadrant of the complex s-plane: $(\rho, \theta) \in \mathbb{R}_{\geq 0} \times \left[0, \frac{\pi}{2}\right]$.

(P2) Finding the Stability Crossing Sets (SCS) between stability and the instability regions. In this case, the searching domain in the complex plane is restraint to $(\rho, \theta) \in \mathbb{R}_{\geq 0} \times \{\frac{\pi}{2}\}$. Hence, only values of $\boldsymbol{\alpha} \in \mathbb{R}^n_{\geq 0}$, for which the poles are crossing the imaginary axis towards the instability region are searched for.

Hence, the problems (P1) and (P2) can be formulated as finding the set of all feasible parameters

$$\boldsymbol{\theta} = (\rho, \theta, \boldsymbol{\alpha})^T \in \Omega = \left(\mathbb{R}_{\geq 0} \times \Theta \times \in \mathbb{R}^n_{\geq 0}\right), \tag{10}$$

where $\Theta = \left[0, \frac{\pi}{2}\right]$ for (P1) and $\Theta = \left\{\frac{\pi}{2}\right\}$ for (P2), satisfying

$$\begin{cases} \mathscr{Re}\{\bar{f}(\rho e^{j\theta}, \boldsymbol{\alpha})\} = 0 \\ \qquad \text{and} \\ \mathscr{Im}\{\bar{f}(\rho e^{j\theta}, \boldsymbol{\alpha})\} = 0 \end{cases} \tag{11}$$

If $\exists \boldsymbol{\theta} = (\rho, \theta, \boldsymbol{\alpha})^T \in \Omega$ such that $\bar{f}(\rho e^{j\theta}, \boldsymbol{\alpha}) = 0$, then, for the problem (P1), the characteristic function has zeros in the closed right half complex plane and, for the problem (P2), on the imaginary axis, which allows determining the SCS.

Both of these problems can be formulated as a Constraint Satisfaction Problem $\mathcal{CSP}$[2]

$$\mathcal{CSP}: \begin{cases} \mathscr{Re}\{\bar{f}(\rho e^{j\theta}, \boldsymbol{\alpha})\} = 0 \\ \mathscr{Im}\{\bar{f}(\rho e^{j\theta}, \boldsymbol{\alpha})\} = 0 \\ 0 < \rho < \mathcal{R}, \quad \theta \in \Theta, \\ \alpha \in \mathbb{R}^n_{\geq 0} \end{cases} \tag{12}$$

where $\mathcal{R}$ is ∞ in theory and is finite in practice for evident implementation reasons. The solution set $\mathbb{S}$ for the problem (12) is rewritten as:

$$\mathbb{S} = \big\{\boldsymbol{\theta} \in \Omega \mid \mathscr{Re}\{\bar{f}(\rho e^{j\theta}, \boldsymbol{\alpha})\} \subset [0] \text{ and } \mathscr{Im}\{\bar{f}(\rho e^{j\theta}, \boldsymbol{\alpha})\} \subset [0]\big\}. \tag{13}$$

The characterization of the whole set $\mathbb{S}$ can be formulated as a set inversion problem:

$$\mathbb{S} = f^{-1}([0]) \cap \Omega, \tag{14}$$

and solved by guaranteed methods using interval arithmetics, introduced in the next subsection.

B. Introduction to interval arithmetics

Interval analysis was initially introduced by Moore [10]. An interval $[x] = [\underline{x}, \overline{x}]$ is a closed, bounded, and connected set of real numbers. The set of all intervals is denoted by $\mathbb{IR}$. Real operations are extended to intervals as follows. Given $[x] \in \mathbb{IR}$ and $[y] \in \mathbb{IR}$:

$$[x] + [y] = [\underline{x} + \underline{y}, \overline{x} + \overline{y}], \tag{15}$$

$$[x] - [y] = [\underline{x} - \overline{y}, \overline{x} - \underline{y}], \tag{16}$$

$$[x] \times [y] = [\min(\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}), \max(\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y})] \tag{17}$$

$$[x]/[y] = \begin{cases} [x] \times \left[\frac{1}{\overline{y}}, \frac{1}{\underline{y}}\right], & \text{if } 0 \notin [y] \\ (-\infty, \infty), & \text{if } 0 \in [y]. \end{cases} \tag{18}$$

Interval arithmetics do not define an algebra because $(\mathbb{IR}, +)$ is not a group. Indeed, elements of $\mathbb{IR}$ do not have an inverse. Take for instance $A = [-1, 1] \in \mathbb{IR}$, then $A + (-A) = [-2, 2]$ is not equal to the degenerated interval $[0] = [0, 0] = \{0\}$. Either, $(\mathbb{IR}, +, *)$ is not a ring etc. Additionally, arithmetic operations on intervals introduce often pessimism because the result of each operation must be included in an interval.

[2] Usually a $\mathcal{CSP}$ is formulated using inequalities

$$\mathcal{CSP}: \begin{cases} \underline{x} \leq \mathscr{Re}\{\bar{f}(\rho e^{j\theta}, \boldsymbol{\alpha})\} \leq \overline{x} \\ \underline{y} \leq \mathscr{Im}\{\bar{f}(\rho e^{j\theta}, \boldsymbol{\alpha})\} \leq \overline{y} \\ 0 < \rho < \mathcal{R}, \quad \theta \in \Theta, \\ \alpha \in \mathbb{R}^n_{\geq 0}, \end{cases}$$

where $\underline{x}, \overline{x}, \underline{y}, \overline{y}$ can also be set to small enough values $-\epsilon, \epsilon, -\epsilon, \epsilon$; as in the fourth initialization of the example in section II-E5.

C. Solving the $\mathcal{CSP}$

1) Contractors: The $\mathcal{CSP}$ (12) is solved by a contractor $\mathcal{C}$, which is an operator which permits to reduce the domain $[\boldsymbol{\theta}]$ without any bisection. Hence, contracting the box $[\boldsymbol{\theta}]$ means replacing it by a smaller box $[\boldsymbol{\theta}]^*$ such that the solution set $\mathbb{S}$ remains unchanged, i.e. $\mathbb{S} \subset [\boldsymbol{\theta}]^* \subset [\boldsymbol{\theta}]$ [11]. There exists different types of contractors depending on whether the system to be solved is linear or not.

In our study, a non linear type contractor named *forward-backward contractor* is used to reduce the initial searching space. The basic idea when implementing this contractor is to decompose a principal constraint into primitive constraints. Each primitive constraint involves elementary operators and functions such as $\{+, -, \times, /, \exp, \log, \ldots\}$. The next example illustrates how a given constraint is used to contract a domain.

2) Example: Consider the constraint:

$$\begin{cases} f(\mathbf{x}) = x_3 - x_2 x_1 = 0, \\ x_1 \in [2, 10],\ x_2 \in [1, 10],\ x_3 \in [1, 5], \end{cases} \tag{19}$$

which can be rewritten as:

$$x_3 = x_2 x_1.$$

The forward interval constraint propagation removes all inconsistent values from $[x_3]$ as follows:

$$[x_3] = ([x_1] \times [x_2]) \cap [x_3] = [2, 5].$$

Then, the backward interval constraint propagation removes all inconsistent values from x_1 and x_2 as follows:

$$[x_1] = ([x_3]/[x_2]) \cap [x_1] = [2, 5],$$
$$[x_2] = ([x_3]/[x_1]) \cap [x_2] = [1, 5/2].$$

After a forward and a backward propagation, the contracted box is $[\mathbf{x}] = \big([2, 5], [1, 5/2], [2, 5]\big)^T$ which contains the solution of the $\mathcal{CSP}$.

In some cases the contractor cannot reduce enough the parameters domain. In such cases, bisection of the variable vector $\boldsymbol{\theta}$ is necessary. The algorithm SIVIA [12], which is described in the following section is based on the association of contractors and splitting.

D. Set Inversion Via Interval Analysis (SIVIA)

This algorithm, proposed by Jaulin and Walter in [12], allows to obtain an inner $\underline{\mathbb{S}}$ and an outer $\overline{\mathbb{S}}$ enclosures of the solution set $\mathbb{S}$ (if it exists), such that:

$$\underline{\mathbb{S}} \subseteq \mathbb{S} \subseteq \overline{\mathbb{S}}. \tag{20}$$

SIVIA is a recursive algorithm based on partitioning of the parameter set into three regions: feasible, undeterminate and unfeasible. SIVIA uses an inclusion test $[t] : \mathbb{IR} \rightarrow \mathbb{N}$ which is a function allowing to prove if an interval $[\boldsymbol{\theta}]$ is feasible in which case it is added to the set $\underline{\mathbb{S}}$. Any undetermined region is bisected and tested again, unless its size $w([\boldsymbol{\theta}])$ is less than a precision parameter η tuned by the user and which ensures that the algorithm terminates after a finite number of iterations. The outer approximation is then computed as $\overline{\mathbb{S}} = \underline{\mathbb{S}} \cup \Delta\mathbb{S}$ where $\Delta\mathbb{S}$ is the union of all remaining undetermined boxes. Hence, the SIVIA algorithm is presented in algorithm 1.

Algorithm SIVIA (in: $[t], [\boldsymbol{\theta}], \eta$; out: $\underline{\mathbb{S}}, \overline{\mathbb{S}}$)

1) Option: Call contractor on $\boldsymbol{\theta}$.
2) If $[t]([\boldsymbol{\theta}]) = [0]$, return;
3) If $[t]([\boldsymbol{\theta}]) = [1]$, then $\underline{\mathbb{S}} := \underline{\mathbb{S}} \cup [\boldsymbol{\theta}]; \overline{\mathbb{S}} := \overline{\mathbb{S}} \cup [\boldsymbol{\theta}]$, return;
4) If $w([\boldsymbol{\theta}]) \leq \eta, \overline{\mathbb{S}} := \overline{\mathbb{S}} \cup [\boldsymbol{\theta}]$;
   Else bisect $[\boldsymbol{\theta}]$ into $[\boldsymbol{\theta}_1]$ and $[\boldsymbol{\theta}_2]$;
5) SIVIA (in: $[t], [\boldsymbol{\theta}_1], \eta$; out: $\underline{\mathbb{S}}, \overline{\mathbb{S}}$);
6) SIVIA (in: $[t], [\boldsymbol{\theta}_2], \eta$; out: $\underline{\mathbb{S}}, \overline{\mathbb{S}}$).

Algorithm 1: The algorithm

The option in line 1 allows either to call the contractor or not at each execution of the SIVIA algorithm which complexity is known to be exponential!

E. Example

Consider the following transfer function having two differentiation orders.

$$F(s, \boldsymbol{\alpha}) = \frac{1}{s^{\alpha_2} + 2s^{\alpha_1} + 1}, \tag{21}$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2) \in \mathcal{A}_1 \times \mathcal{A}_2 \subset \mathbb{R}^2_{\geq 0}$, $\mathcal{A}_1$ and $\mathcal{A}_2$ define the searching domains. It can be analyzed by checking the position of the zeros of the characteristic function

$$f(s, \boldsymbol{\alpha}) = s^{\alpha_2} + 2s^{\alpha_1} + 1 \tag{22}$$

$$\bar{f}(\rho e^{j\theta}, \boldsymbol{\alpha}) = \rho^{\alpha_2}\cos(\theta\alpha_2) + 2\rho^{\alpha_1}\cos(\theta\alpha_1) + 1 + \mathrm{j}\left(\rho^{\alpha_2}\sin(\theta\alpha_2) + 2\rho^{\alpha_1}\sin(\theta\alpha_1)\right) \tag{23}$$

1) Implementing the forward-backward contractor on the system under study: A first contractor could be implemented, after the real part of $\bar{f}$:

$$\mathscr{Re}\{\bar{f}(\rho e^{j\theta}, \boldsymbol{\alpha})\} = 0 \Leftrightarrow \rho^{\alpha_2}\cos(\theta\alpha_2) + 1 = -2\rho^{\alpha_1}\cos(\theta\alpha_1) \tag{24}$$

A second one could also be implemented, after the imaginary part of $\bar{f}$:

$$\mathscr{Im}\{\bar{f}(\rho e^{j\theta}, \boldsymbol{\alpha})\} = 0 \Leftrightarrow \rho^{\alpha_2}\sin(\theta\alpha_2) = -2\rho^{\alpha_1}\sin(\theta\alpha_1) \tag{25}$$

However, handling sin and cos functions in each contractor is not an easy task because asin and acos functions return angles in their principal determination, i.e. between 0 and π for the acos, and between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$ for the asin. In that case, care must be taken to set back the angles to the correct determination. Another alternative, is to combine (24) and (25) to obtain another contractor with less sin and cos functions. Such a contractor, named *combined contractor*, is obtained by squaring both equations and summing them up

$$\rho^{2\alpha_2} + 2\rho^{\alpha_2}\cos(\theta\alpha_2) + 1 = 4\rho^{2\alpha_1}. \tag{26}$$

```
function [x]=Comb_Contractor_Red(x)
 global nb_siv;
 xx = x;
 rho = x(1); theta = x(2);
 alpha1 = x(3); alpha2 = x(4);
%Forward
 x1 = rho^(2*alpha2);
 x2 = 2*rho^alpha2;
 x3 = theta*alpha2;
 x4 = cos(x3);
 x5 = x2*x4;
 x6 = x1 + x5 + 1;
 x7 = 4*rho^(2*alpha1);
%Backward
 x7     = intersect(x6, x7);
 alpha1= intersect(alpha1,1/2*log(x7/4)/log(rho));
 rho    = intersect(rho, (x7/4)^(1/(2*alpha1)));
 x6     = intersect(x6, x7);
 x5     = intersect(x5, x6 - x1 - 1);
 x2     = intersect(x2, x5/x4);
 alpha2= intersect(alpha2, log(x2/2)/log(rho));
 rho    = intersect(rho, (x2/2)^(1/alpha2));
 x1     = intersect(x1, x6 - x5 - 1);
 rho    = intersect(rho, x1^(1/(2*alpha2)));
 alpha2= intersect(alpha2, log(x1)/(2*log(rho)));

 x = [rho, theta, alpha1, alpha2];
 if any(isnan(x))
     x=xx;
 end
end
```

Fig. 1. The implementation of the combined contractor (26) using the IntLab toolbox [13] under Matlab.

A single cos function remains in (26) instead of two in the previous two contractors, which is easier to handle. This contractor is implemented in Fig.1, using the IntLab toolbox [13] under Matlab.

The algorithm is applied to the characteristic function (22), using four different initializations. In the first three, the problem (P2) is considered and in the fourth, the problem (P1) is treated.

2) First initialization: The initial searching box and tolerance are respectively set to:

$$\boldsymbol{\theta} = (\rho, \theta, \alpha_1, \alpha_2)^T \in [0,4] \times \left\{\frac{\pi}{2}\right\} \times [0,3] \times [0,4.5] \quad (27)$$

$$\eta = \mathrm{diam}(\boldsymbol{\theta})/2^7 \quad (28)$$

where $\mathrm{diam}(\boldsymbol{\theta})$ defines the length of each element of $(\boldsymbol{\theta})$.

The SIVIA algorithm is executed:

- without contractors (without step 1 in the algorithm). In this case the SIVIA function is called 13 543 times in 190 sec. The obtained outer enclosure $\overline{\mathbb{S}}$ is plotted in Fig.2.
- with the combined contractor (26) called at each step of the SIVIA algorithm (with step 1 in the algorithm). The SIVIA function is called 8 711 times in 296 sec. The obtained outer enclosure $\overline{\mathbb{S}}$ is plotted in Fig.3.

Moreover, the values at which the poles cross the imaginary axis correspond more or less exactly in both cased to the plot of Fig.4, which validates *a posteriori* that all the poles are inside the searching interval $\rho \in [0, 4]$. In case some poles were touching the limit $\mathcal{R} = 4$, it would have been necessary to choose a bigger $\mathcal{R}$.

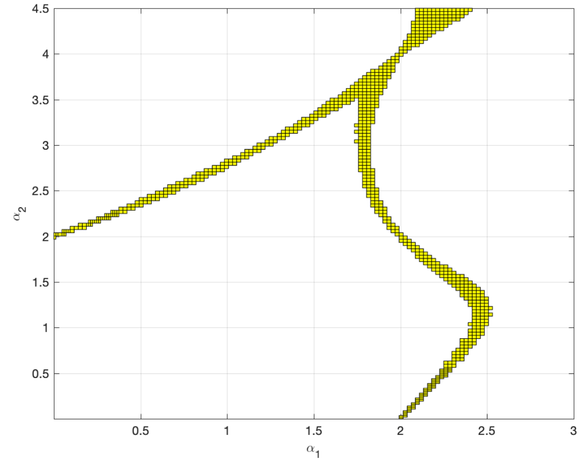


Fig. 2. First initialisation – Stability crossing sets obtained **without contractors**. Zeros of the characteristic function $\tilde{f}$ which arguments equal $\frac{\pi}{2}$ are probably contained in the yellow boundary (outer enclosure $\overline{\mathbb{S}}$). The lower left region delimited by the yellow boundary represents the guaranteed stability region.

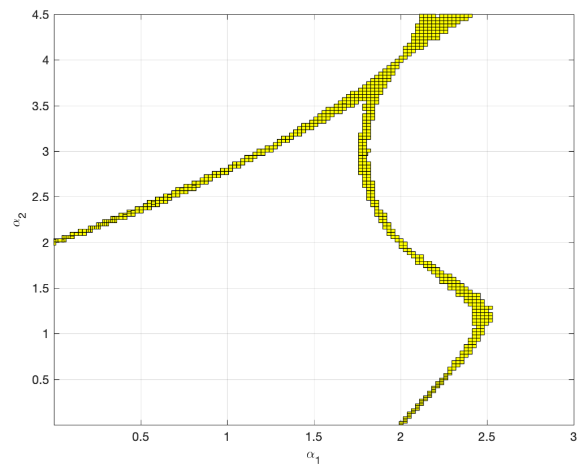


Fig. 3. The same as Fig.2, however **with contractors**.

As a conclusion, regarding this first initialization, the algorithm using the combined contractor is a little bit more precise for the same tolerance factor η. Execution speeds of both algorithms are comparable. The former as compared to the latter converges in a bigger number of iterations, however quicker, because the latter calls the contractor at each SIVIA iteration.

3) Second initialization: Let's search for the SCS by enlarging the searching domain. The initial box is now set to:

$$\boldsymbol{\theta} = (\rho, \theta, \alpha_1, \alpha_2)^T \in [0,4] \times \left\{\frac{\pi}{2}\right\} \times [0,15] \times [0,20] \quad (29)$$

The tolerance is defined as in (28), however applied to the new definition of the initial searching box.

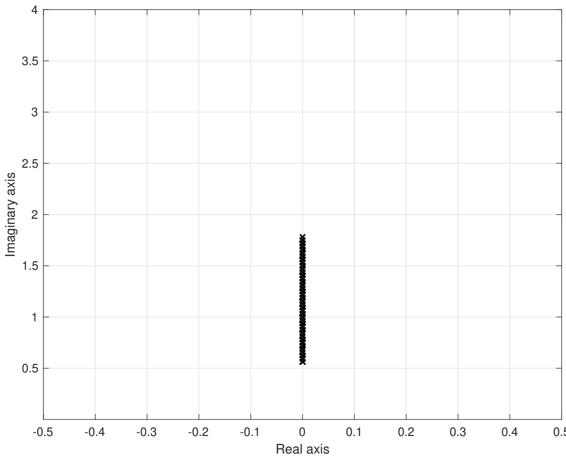


Fig. 4. First initialization – Zeros of the characteristic function $\bar{f}$ crossing the imaginary axis.

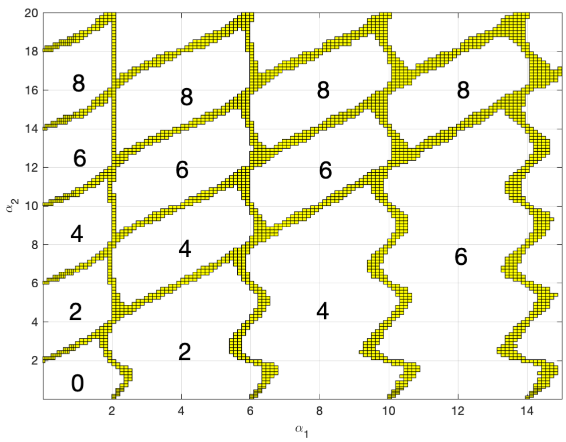


Fig. 5. Second initialization – Stability Crossing Sets (wider intervals as compared to Figs.2 and 3). The number of unstable poles is indicated in each region.

The SIVIA algorithm, without contractors, is called 73 037 times in 1 090 seconds. The obtained outer enclosure $\overline{\mathbb{S}}$ of the SCS is plotted in Figs 5, which indicates additionally the number of unstable poles, computed at integer values of the parametric points. Hence, all systems with parameters inside the different regions delimited by the SCS have the indicated number of poles. The intervals of poles look very much like the ones in Fig.4.

4) Third initialization: A major change is operated here. Instead of testing, the $\mathcal{CSP}$ defined in (12), a new $\mathcal{CSPN}$ is defined by enlarging the acceptable mapping of $\bar{f}(\rho e^{j\theta}, \boldsymbol{\alpha})$ to a square of size ϵ instead of a single point (the origine).

$$\mathcal{CSPN}\colon \begin{cases} -\epsilon \leq \mathscr{Re}\{\bar{f}(\rho e^{j\theta}, \boldsymbol{\alpha})\} \leq \epsilon, \\ -\epsilon \leq \mathscr{Im}\{\bar{f}(\rho e^{j\theta}, \boldsymbol{\alpha})\} \leq \epsilon, \\ 0 < \rho < \infty, \quad \theta \in \Theta, \\ 0 < \alpha_1 < \infty, \; 0 < \alpha_2 < \infty, \\ \epsilon = 0.1 \end{cases} \tag{30}$$

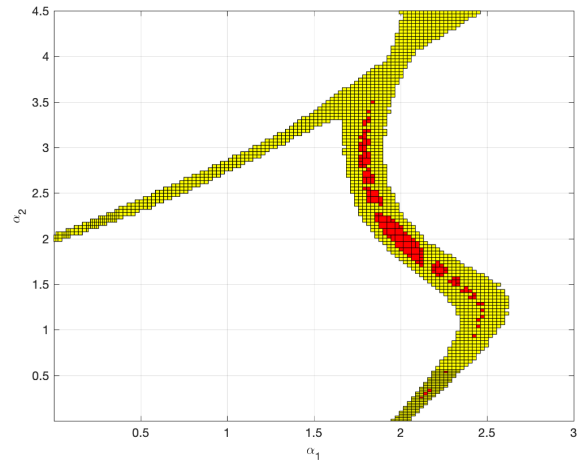


Fig. 6. Third initialization – Inner $\underline{\mathbb{S}}$ (in red), and Outer $\overline{\mathbb{S}}$ (in yellow) enclosures of the $\mathcal{CSPN}$ defined in (30)

Hence, instead of searching for the zeros of the characteristic function $\bar{f}$ in (22), the algorithm searches for intervals $[\boldsymbol{\theta}]$ that are mapped according to $\bar{f}$ inside a square of length ϵ. This is the usual way $\mathcal{CSP}$s are formulated. The same parameters and tolerance are chosen as in the first initialization in (27) and (28).

The results, obtained without contractors in 27 795 iterations and 443 sec, are plotted in Fig.9, where red and yellow parts indicate the inner and the outer enclosures $\underline{\mathbb{S}}$ and $\overline{\mathbb{S}}$ of (20).

It turns out not to be interesting to consider the $\mathcal{CSPN}$ (30) instead of the initial $\mathcal{CSP}$ (12), because it widens the feasible solution set as the square, of length ϵ, defining the admissible mapping gets wider.

5) Fourth initialization: In this part, the problem (P1) is solved. Hence, instead of looking for the stability crossing sets, let's look for all the zeros of $\bar{f}(\rho e^{j\theta}, \boldsymbol{\alpha})$ in the first quadrant. Consider the $\mathcal{CSP}$ in (12), and the following searching box:

$$\boldsymbol{\theta} = (\rho, \theta, \alpha_1, \alpha_2)^T \in [0, 4] \times \left[0, \frac{\pi}{2}\right] \times [0, 3] \times [0, 4.5] \tag{31}$$

When setting the tolerance to (28), the algorithm is stopped after an hour because of convergence issues. Then, the tolerance is augmented to:

$$\eta = \mathrm{diam}(\boldsymbol{\theta})/2^4$$

The algorithm converges in 12 525 iterations and 194 sec. The obtained outer enclosure $\overline{\mathbb{S}}$ is plotted in Fig.7.

Apparently, the root-searching-domain in the first quadrant, is validated *a posteriori* in Fig.8: all the poles of the first quadrant are inside the searching domain, defined by $\rho \in [0, 4]$, when $(\alpha_1, \alpha_2) \in \times[0, 3] \times [0, 4.5]$.

Higher precision is definitely required to find out a better sketch of the stability region (in white).

However, this problem appears to be ill-posed as the $\mathcal{CSP}$ (12) evaluated for interval values of $[\boldsymbol{\theta}]$, can never be satisfied. A mapping of $[\boldsymbol{\theta}]$ with $\bar{f}$ is an interval that can never be a subset of $\{0\}$. Hence, in the instability region, the algorithm

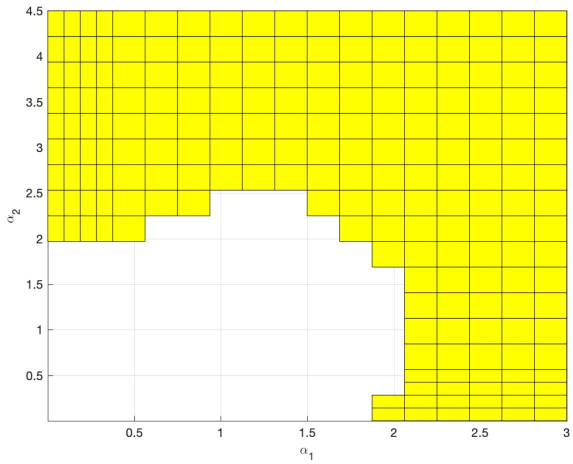


Fig. 7. Fourth initialization – Guaranteed stability in white and possible instability (outer enclosure $\overline{\mathbb{S}}$) in yellow

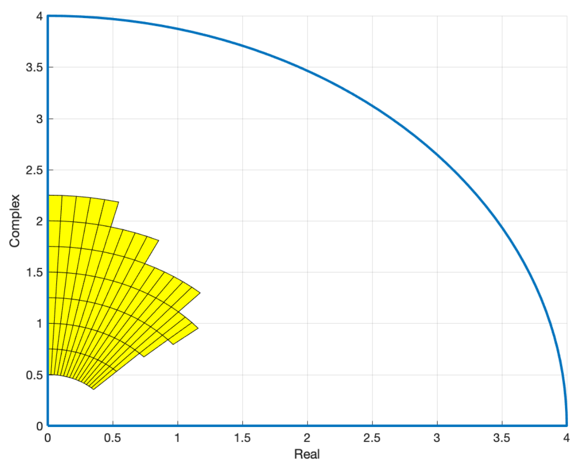


Fig. 8. Fourth initialization – Possible root location in yellow, searching domain boundary in blue

will keep bisecting, until reaching the precision η. It turns out that the time complexity of the SIVIA algorithm is higher than a brute-force search on boxes of elementary sizes η, which is not interesting.

As a conclusion of this part, it turns out that it is more interesting to solve the problem (P2) by looking for the stability crossing sets and deducing the stability regions.

III. Which norm for fractional systems?

This section of the paper contains results originally published in [2].

As mentioned previously, (4) is only a necessary condition. The system might be $\mathscr{L}_p$-stable and yet have an impulse response, f, with an infinite $\mathscr{L}_p$-norm $\forall p \in [1, \infty]$. The following theorem, proven in [2], states the existence conditions of the $\mathscr{L}_p$-norms.

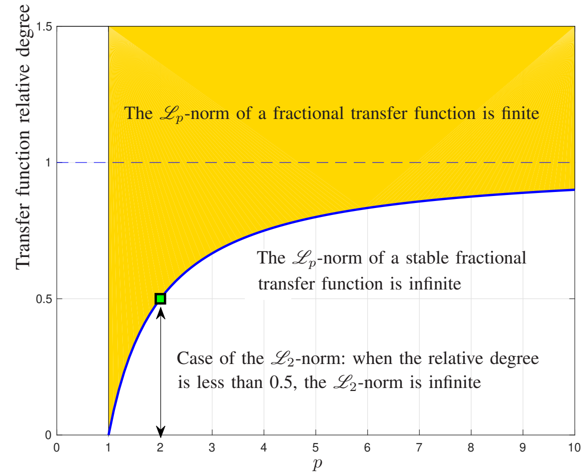


Fig. 9. Finiteness region of the $\mathscr{L}_p$-norm of fractional stable transfer functions in the relative-degree versus p plane. The border curve $(1 - \frac{1}{p})$ does not belong to the finiteness region.

Theorem 2. *Let a fractional transfer function as*

$$\tilde{F}(s) = \frac{F(s)}{s^{\mu}} \tag{32}$$

where $F(s)$, given by (1), is BIBO-stable and where $\mu \geq 0$. Numerator and denominator of $F(s)$ are assumed to be coprime (with no possible simplification between poles and zeros). Then, the $\mathscr{L}_p$-norm, $1 \leq p \leq \infty$, of the impulse response of $\tilde{F}(s)$ is finite if and only if the transfer function relative degree satisfies:

$$\mu + \alpha_N - \beta_M > 1 - \frac{1}{p} \tag{33}$$

and the integrator order satisfies:

$$0 < \mu < 1 - \frac{1}{p} \tag{34}$$

or

$$\mu = 0 \tag{35}$$

The yellow zone in Fig.9 shows finite combinations of $\mathscr{L}_p$-norms, in the plane relative-degree versus p. Similarly, the orange zone in Fig.10 shows finite combinations of $\mathscr{L}_p$-norms, in the plane integrator-order versus p.

Remarks:

- $\mathscr{L}_p$-norm finiteness conditions (33)-(34) are in accordance with the $\mathscr{L}_2$-norm finiteness conditions determined in [14].
- Equation (33) shows that all the $\mathscr{L}_p$-norms of rational systems, $\forall 1 \leq p \leq \infty$, are always finite because the relative degree is an integer at least equal to one (for a proper transfer function with no nonzero feedthrough gain). Additionally, (34) shows that the $\mathscr{L}_p$-norms, $\forall 1 \leq p \leq \infty$, are always infinite in presence of a rational integrator, with $\mu = 1$.
- A pure integrator $\frac{1}{s^{\mu}}$, $\forall \mu \in \mathbb{R}_{>0}$, has always an infinite $\mathscr{L}_p$-norm, $\forall 1 \leq p \leq \infty$, because conditions $\mu + \alpha_N -$

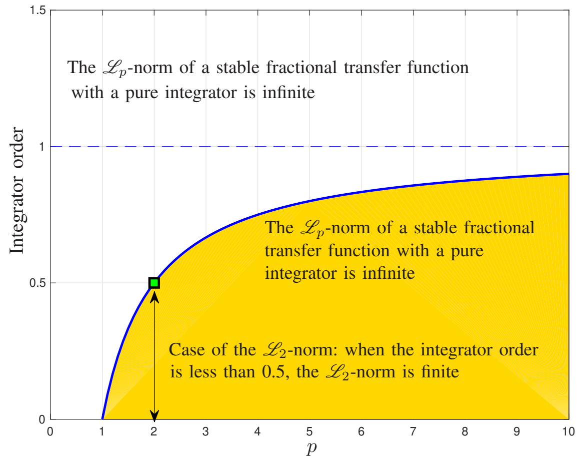


Fig. 10. Finiteness region of the $\mathscr{L}_p$-norm of fractional stable transfer functions in the integrator-order versus p plane. The border curve $(1 - \frac{1}{p})$ does not belong to the finiteness region, except the point $(p = 1, \mu = 0)$ which does.

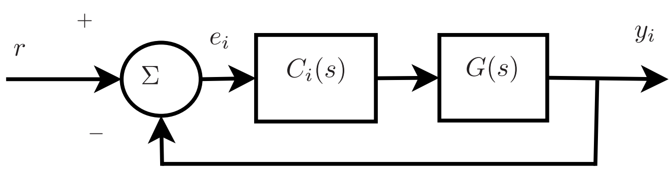


Fig. 11. A simple feedback control system structure

$\beta_M > (1 - \frac{1}{p})$ and $\mu < (1 - \frac{1}{p})$, cannot be satisfied simultaneously (here $\alpha_N = \beta_M = 0$) .

In the following example, the results are used to choose a proper criterion for evaluating performance of a feedback control loop.

IV. Example

This example is taken from [15, example 2]. Consider the simple feedback control system structure of Fig.11 and different fractional order PI controllers, with $i \in \{1, 2, 3, 4\}$,

$$C_i(s) = K_{p_i} + \frac{K_{I_i}}{s^{\lambda_i}}, \tag{36}$$

which yield different closed-loop transfer functions:

$$T_i(s) = \frac{C_i(s)G_i(s)}{1 + C_i(s)G_i(s)} \tag{37}$$

as reported in Table 1. Tavazoei (2010) evaluates numerically different integral performance indices among which the integral of absolute error (IAE) and the integral of squared error (ISE), on a step response. The IAE and ISE-indices are respectively the $\mathscr{L}_1$-norm and the $\mathscr{L}_2$-norm squared of the error signal $e_i(t)$ of Fig.11 (respectively $\|e_i\|_1$ and $\|e_i\|_2^2$) when the input $r(t)$ is a step:

$$E_i(s) = \frac{1}{s}(1 - T_i(s)) \tag{38}$$

The $\mathscr{L}_2$-norm of $e_i(t)$ was also computed analytically in [14] and allowed to confirm the results announced in [15] regarding the ISE-index. In this paper, the $\mathscr{L}_4$-norm of $e_i(s)$ is additionally computed by numerical integration of the time-domain signals $e_i(t)$. The $\mathscr{L}_\infty$-norm is deduced easily.

Note that $E_1(s)$ has a proper integrator of order 0.2 and hence, according to (34), an infinite $\mathscr{L}_1$-norm and finite $\mathscr{L}_p$-norms for $p > 1.25$. Consequently, $\|e_i\|_2$, $\|e_i\|_4$, and $\|e_i\|_\infty$ are finite. $E_4(s)$ has a proper integrator of order 0.5 and hence infinite $\mathscr{L}_1$ and $\mathscr{L}_2$-norms. Additionally, [15] has evaluated other performance indices such as the *integral of time multiplied absolute error* (ITAE), the *integral of time multiplied squared error* (ITSE), and the *integral of squared of time multiplied error* (ITSE). All these performance indices were shown to be infinite for $E_4(s)$. No finite performance index has been proposed in [15] for evaluating the output feedback control law for $E_4(s)$. Theorem 2 and condition (34) show that the $\mathscr{L}_p$-norm is finite for all $p > 2$. Here, the $\mathscr{L}_4$-norm of the error signal can be used as a finite performance index of the output feedback control. For the remaining systems, $E_2(s)$ and $E_3(s)$ have no integrators, relative degrees greater than $(1 - \frac{1}{p})$, stable $s^{0.5}$-poles, and hence finite $\mathscr{L}_p$-norms $\forall 1 \leq p \leq \infty$. Note that $\|e_i\|_\infty$ equals 1 for all i, because the step response always starts at $y_i(0) = 0$ and hence $\|e_i\|_\infty = e_i(0) = 1$. Consequently, the $\mathscr{L}_\infty$-norm is not, in this case, an interesting performance index.

V. Conclusions

This paper proposes an algorithm, based on interval arithmetics, for stability analysis of fractional transfer functions. Guaranteed stability region is determined in the parametric space. Two problems have been formulated and it has been shown that the problem of finding the parametric region for which the system is unstable is ill-posed because the bisection algorithm has a time-complexity worse than a brute-force search. However, the problem of finding stability crossing sets turns out to be very interesting, as it allows finding with a reasonable complexity, the stability crossing sets and hence deducing the whole stability region.

Having stable fractional transfer functions does not, however guarantee the existence of the $\mathscr{L}_p$-norms of its impulse response. Some additional conditions on its relative degree must be fulfilled. This helps choosing a performance criterion for feedback control loops.

The analytical computation of the $\mathscr{L}_2$-norm was proposed in [14] for commensurate systems only. A challanging task would be to extend this result to incommensurate systems.

References

[1] R. Malti, M. R. Rapaić, and V. Turkulov, "Stability analysis of incommensurate elementary fractional systems using interval arithmetics," in *International Conference on Fractional Differentiation and its Applications (ICFDA)*, Warsaw, Poland, 2021.

[2] R. Malti, "A note on L_p-norms of fractional systems," *Automatica*, vol. 49, no. 9, pp. 2923–2927, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0005109813003166

[3] K. Oldham and J. Spanier, "The replacement of Fick's laws by a formulation involving semi-differentiation," *Electro-anal. Chem. Interfacial Electrochem*, vol. 26, pp. 331–341, 1970.

[4] ——, *The fractionnal calculus - Theory and Applications of Differentiation and Integration to Arbitrary Order*. Academic Press, New-York and London, 1974.

i	$T_i(s)$	$E_i(s)$	$\|e_i\|_1$-IAE	$\|e_i\|_2^2$-ISE	$\|e_i\|_4^4$	$\|e_i\|_\infty$
1	$\frac{2s^{0.8}+1}{2s^{2.4}+s^{1.6}+s^{0.8}+1}$	$\frac{2s^{1.6}+s^{0.8}-1}{2s^{2.6}+s^{1.8}+s+s^{0.2}}$	∞	3.43	2.28	1
2	$\frac{2s^{1.6}+s^{0.8}+1}{2s^{2.4}+s^{1.6}+s^{0.8}+1}$	$\frac{2s^{1.4}-s^{0.6}}{2s^{2.4}+s^{1.6}+s^{0.8}+1}$	3.48	1.10	0.27	1
3	$\frac{s^{1.6}+s^{0.8}+1}{2s^{2.4}+s^{1.6}+s^{0.8}+1}$	$\frac{2s^{1.4}}{2s^{2.4}+s^{1.6}+s^{0.8}+1}$	3.32	1.11	0.35	1
4	$\frac{2s^{0.5}+1}{2s^{1.5}+s+s^{0.5}+1}$	$\frac{2s+s^{0.5}-1}{2s^{2}+s^{1.5}+s+s^{0.5}}$	∞	∞	0.27	1

TABLE I

$\mathscr{L}_1, \mathscr{L}_2, \mathscr{L}_4$, AND $\mathscr{L}_\infty$,-NORMS OF $e_i(t)$ FOR DIFFERENT OUTPUT FEEDBACK CONTROL SYSTEMS TAKEN FROM (TAVAZOEI, 2010, EXAMPLE 2).

[5] D. Matignon, “Stability properties for generalized fractional differential systems,” *ESAIM proceedings - Systèmes Différentiels Fractionnaires - Modèles, Méthodes et Applications*, vol. 5, 1998.

[6] R. Malti, X. Moreau, F. Khemane, and A. Oustaloup, “Stability and resonance conditions of elementary fractional transfer functions,” *Automatica*, vol. 47, no. 11, pp. 2462–2467, 2011.

[7] C. Hwang and Y.-C. Cheng, “A numerical algorithm for stability testing of fractional delay systems,” *Automatica*, vol. 42, no. 5, pp. 825 – 831, 2006.

[8] J. Trigeassou, A. Benchellal, N. Maamri, and T. Poinot, “A frequency approach to the stability of fractional differential equations,” *Transactions on Systems, Signals & Devices (TSSD)*, vol. 4, no. 1, pp. 1 – 25, 2009.

[9] C. Bonnet and J. Partington, “Coprime factorizations and stability of fractional differential systems,” *Systems & Control Letters*, vol. 41, no. 3, pp. 167 – 174, 2000.

[10] R. Moore, *Interval analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1966.

[11] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter, *Applied interval analysis*. London: Springer-Verlag, 2001.

[12] L. Jaulin and E. Walter, “Set inversion via interval analysis for nonlinear bounded-error estimation,” *Automatica*, vol. 29, no. 4, pp. 1053–1064, 1993.

[13] S. Rump, *Developments in reliable computing*. Kluwer academic publisher, 1999, ch. IntLab – Interval laboratory, pp. 77–104.

[14] R. Malti, M. Aoun, F. Levron, and A. Oustaloup, “Analytical computation of the H_2-norm of fractional commensurate transfer functions,” *Automatica*, vol. 47, no. 11, pp. 2425–2432, 2011.

[15] M. Tavazoei, “Notes on integral performance indices in fractional-order control systems,” *Journal of Process Control*, vol. 20, no. 3, pp. 285 – 291, 2010. [Online]. Available: http://www.sciencedirect.com/science/article/B6V4N-4XHC6CK-1/2/20d020ac73210cb2ce6f0f6ab89e92e2

Application of cascade control in the process of flue gas desulfurization of thermal power plant

Goran Kvascev, Zeljko Djurovic, Avram Avramovic

***Abstract*—The paper presents the use of cascade control for the needs of an efficient flue-gas desulphurization process in the thermal power plant TE KO Drmno. A system for reducing the content of sulfur oxides (SO2) in flue exhaust gases - desulphurization has been implemented within two thermal power plants units. The technological process, control structure, implementation of cascade control, and plant operation results are presented. By applying the proposed control structure, the efficient operation is achieved of the entire system in terms of the regulations of the European Commission in terms of emissions of sulfur oxides and particles, but also electricity consumption and energy efficiency. The goals that the plant was supposed to meet were achieved, as well as the economic justification of the entire project.**

***Index Terms*— Flue-gas desulphurization, thermal power plant, cascade control**

I. Introduction

In the Republic of Serbia, within the PE “Electric Power Industry of Serbia” (EPS), 34,896 GWh of electricity is produced annually, based on ten-year average. The electricity generation capacities owned by EPS have a total capacity of 7,855 MW and consist of 22 thermoblocks, 49 hydro units, and one reversible hydro power plant with 2 units. About 70% of production comes from thermal power plants and about 30% from hydropower plants. As fuel in thermal power plants, coal from surface mines is mostly used, i.e. lignite, whose average annual production ranges from 37 to 40 million tons of coal.

Lignite is a solid fuel that contains a high percentage of sulfur. Combustion of sulfur-containing fuel produces sulfur dioxide SO_2, as a dominant product of its oxidation, and sulfur trioxide SO_3 (in the amount of several percent of SO_2), as well as other oxides sulfur, which have no greater significance. Sulfur oxides SO_2 and SO_3 have been recognized as the most common and most dangerous gases of anthropogenic origin, with a serious negative impact on human health and vegetation, primarily creating acid rain. Therefore, it is of special interest to reduce the emission limit values (ELVs) of sulfur oxides to an acceptable level, which will not be harmful to the environment and the health of the population.

Goran Kvaščev is with the University of Belgrade, School of Electrical Engineering, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mail: kvascev@etf.bg.ac.rs)
Željko Đurović is with the University of Belgrade, School of Electrical Engineering, , Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mail: zdjurovic@etf.bg.ac.rs)
Avram Avramovic is with the TE DRMNO - Kostolac, Serbia (e-mail: avram.avramovic@te-ko.rs)

In the Republic of Serbia, based on our conditions, ELVs are defined by several criteria:

- whether plants are fired by solid, liquid, or gaseous fuels,
- whether plants are small, medium, or large (in terms of power),
- whether plants are old or new.

In the Republic of Serbia, based on: “Uredba o graničnim vrednostima emisija zagađujućih materija u vazduhu iz postrojenja za sagorevanje” (“Sl. glasnik RS”, br. 6/2016) [1], the emission limit values, ELVs, for sulfur dioxide are expressed in mg/normal m^3 as function of power of plant, and which are applied to old plants, are given in the Figure 1.

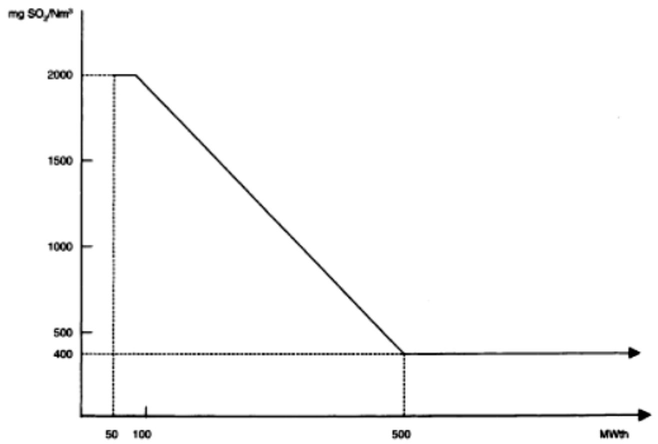


Fig. 1. Emission limit values, ELVs, for sulfur dioxide in mg/normal m3, which are applied to old plants as function of power of power plant

For plants with a thermal power of 100 to 500 MWth, the emission limit values for sulfur dioxide are calculated according to:

$$y = -4x + 2400$$

where: x - thermal power of the plant (MWth), y - emission limit value for SO_2 (mg SO_2/normal m^3).

For plants with a thermal power of more than 500 MWth, a desulphurization rate of at least 94% must be achieved.

On the other hand, observing the *National Plan for the Reduction of Emissions of Major Pollutants from Old Large Combustion Plants* [2] (*Nacionalni plan za smanjenje emisija* *glavnih zagađujućih materija koje potiču iz starih velikih postrojenja za sagorevanje*), which clearly defines the maximum emissions of sulfur oxides, it is clear that these goals have not yet been achieved. If we take as an example the TE KO Drmno B1/B2 units, for which the maximum emission of 7,957.03 tons per year is defined according to the plan [2], and compare with the data that over 95,000 tons of SO_2 were emitted during the previous year, it is clear that it is necessary to provide a plant for processing sulfur oxides, i.e. flue-gas desulphurization system. Such a system has been put into operation during the past period. Tests, final adjustments of control loops, and obtaining the necessary approvals for work are in progress.

II. FLUE-GAS DESULPHURIZATION TE KO DRMNO

According to the conclusions of the Study "*Directions of optimal reduction of sulfur oxide emissions from thermal power plants of the Electric Power Industry of Serbia*" [3], TE KO Drmo B1/B2 was chosen as the first thermal power plant where the construction of such a plant is planned. The flue gas desulphurization plant in TE KO Drmno is the first desulphurization system in Serbia to be operational, and it started operating on October 23, 2020. Based on previously accepted analyzes, the wet limestone / gypsum process was defined as the reference desulphurization process.

Process Overview

The process of flue gas desulfurization consists of several chemical processes that are essential for understanding and applying the control system, which is presented in Figure 2:

- The SO_2 in the flue gas is absorbed into the circulating slurry
- Circulating slurry containing calcium sulfate ($CaSO_4$) and precipitated limestone
- Add some limestone into absorb to supply alkalinity and calcium ions (Ca ++) source
- Add air into absorb for Oxidized SO_2 to sulphate (SO_4).
- Solution sulfate ion and calcium ion reaction precipitated gypsum ($CaSO_4 * 2H_2O$)
- The precipitated gypsum is separated from the slurry water is removed. This results in solid by-product from the process. Slurry liquid return.

The process schematic is shown in Figure 3, which shows the thermal power plant that was upgraded by the flue-gas desulphurization subsystem, and which consists of three main parts:

- Absorber – within which the process of Absorption, Oxidation, and Neutralization (shown in Figure 2) takes place
- Absorbers preparation
- Gypsum dehydration

The most important process takes place in the Absorber, where the main control structures are located. It is necessary to control the flow of an absorber with the control valve in order to make chemical reaction to be successful and efficient. This can be achieved by adding limestone into the absorber to ensure alkalinity. Constant alkalinity produces a successful and efficient flue-gas desulphurization process. As the absorber system is large in volume, and the time constants are also large (on the order of 15 minutes), it is extremely important to accurately add limestone slurry. This is main task for the control system, to work efficiently and precisely. For this control structure, cascade control is proposed, with the outer loop according to alkalinity (with reference pH value), and an inner loop for control of a flow of limestone slurry.

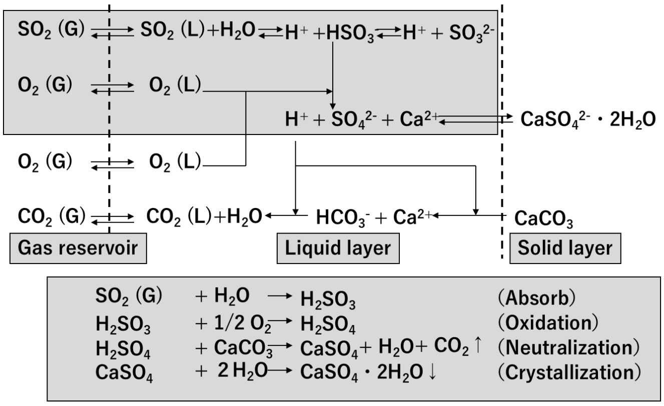


Fig. 2. Chemical Process Overview

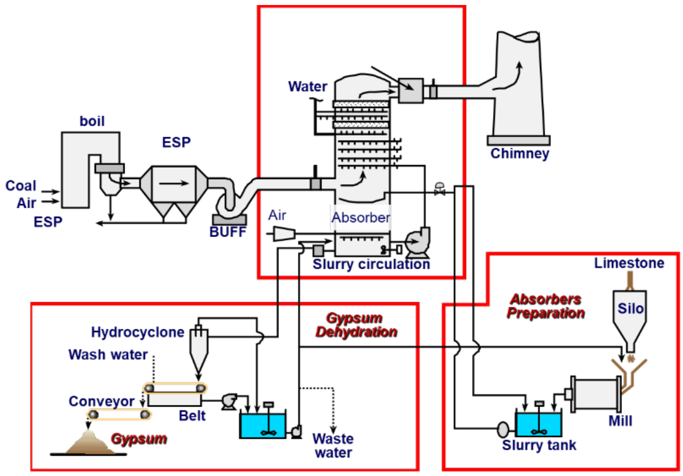


Fig. 3 FGD Process schematic

Figure 4 shows a process flow diagram (PFD):

- flue gases: when the flue gas system is running: VDG → ODG inlet valves → Buster fan → Absorber → New chimney
- flue gases: when the flue gas system is not working: VDG → Bypass valves → Old chimney
- air: when the air sealing system is running: Air sealing fan → Heater → Valve closed
- process water: when the process of flushing the absorber inlet with water in case of emergency is in progress: Process water tank → Flushing water pump in case of emergency → Flushing water tank → Flue gas duct spraying

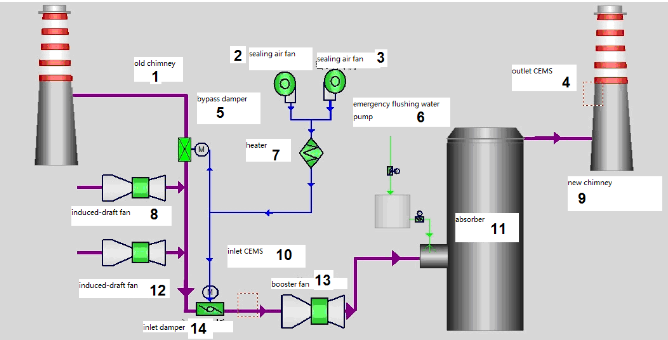


Fig. 4 Process flow diagram, 1-old chimney, 2,3-sealing fans, 4-outlet CEMS, 5-bypass damper, 6-emergency flushing water pump, 7-heater, 8,12-induced-draft fans, 9-new chimney, 10-inlet CEMS, 11-absorber, 13-booster fan, 14-inlet damper

III. Control Loops

Figure 5 shows the simplified P&I diagram shown on DCS SCADA system - Absorber picture. Measurements, state of actuators (valves, pumps, and actuators), as well as controls within control loops are presented. Control valve (measured position: 2TJ39S001) is responsible for alkalinity (pH measurement marked as 2TD40A001 and 2TD40A002). The flow of liquid limestone slurry through this pipe is marked as 2TJ39F001. As the system for measuring the pH value needs to be maintained at regular time intervals, a double measurement is planned, so the value of one probe is taken as main and another as spare, which gives a regular measurement in full time.

Within this research, a cascade control structure was proposed, with the use of PI / PID controllers, shown in Figure 6. Simpler control procedures did not give satisfactory results, observing a testing period of 6 months.

Cascade control structure consists of two loops:

- The outer loop controls the alkalinity (pH value) of the PID controller - the pH value is controlled by setting the flow of the internal PI controller. Alkalinity measurements is achieved by sensors
- Internal PI control loop - controls the reference flow from the outer loop through control valve 2TJ39S001, and measured flow of limestone slurry – 2TJ39F001

PI/PID controllers are tuned using the SRT method [4], measuring the response in the open loop in the nominal mode of plant. In the first version, the control of the PID controller in the external loop was set without differential action, however, after testing, it was necessary to introduce significant differential action for two reasons: faster elimination of disturbances and problems due to jamming of the control valve due to longer periods without moving the actuator.

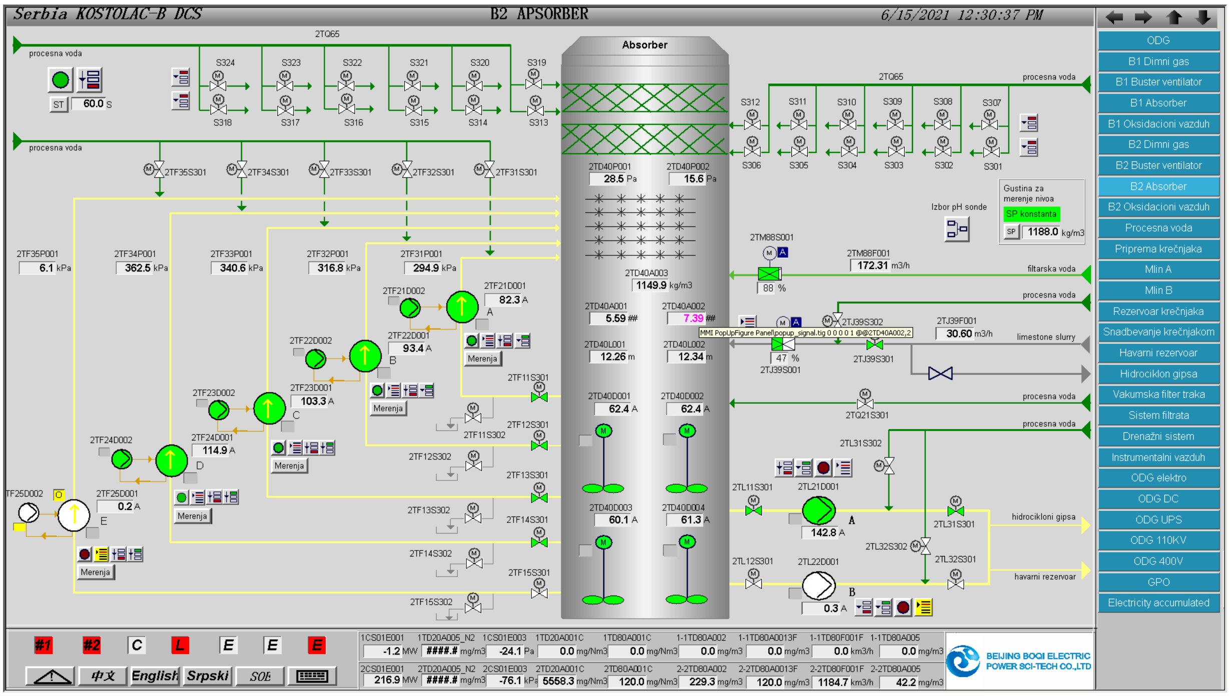


Fig. 5 SCADA control system – P&I diagram

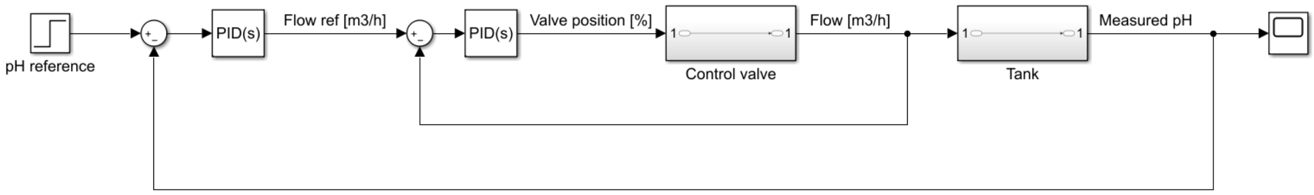


Fig. 6 Proposed cascade control structure

IV. Main Results

Figures 7 and 8 show the measurements from the plant after the final testing in the nominal regime. Figure 7 shows satisfactory results of control the alkalinity (pH value) with a rise time of about 10 minutes with a step change of the reference value. The measured flow in inner loop is almost same as flow reference.

Observing the measured values in the real environment, it can be concluded that the control of the plant operates within the specified standards. The emission of sulfur oxides SO_2 does not exceed the value of 250 mg/Nm^3, measured in 24 hours time range. Considering that the inlet concentration of SO_2 is in the range of about 5000mg/Nm^3, it can be confirmed that the efficiency of the desulfurization process is more then 99%.

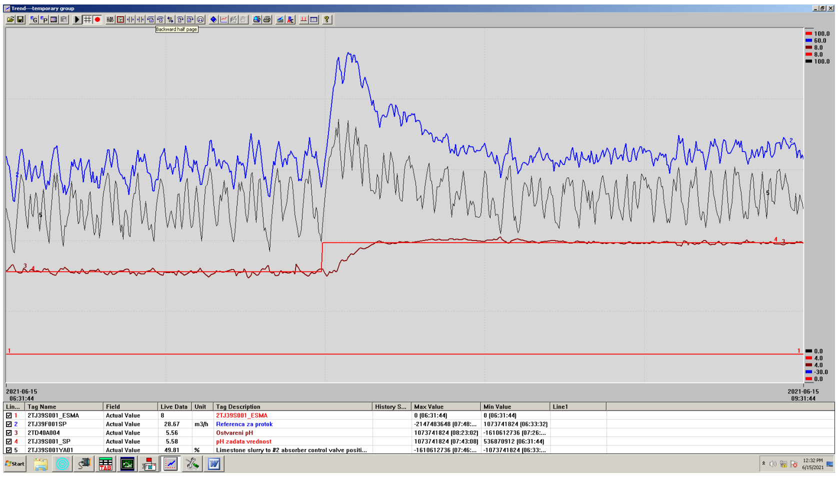


Fig. 7. Control performance pH control loops - 3h time scale: pH reference value – red color, pH measured value – brown color, flow reference – blue color, valve position – gray color

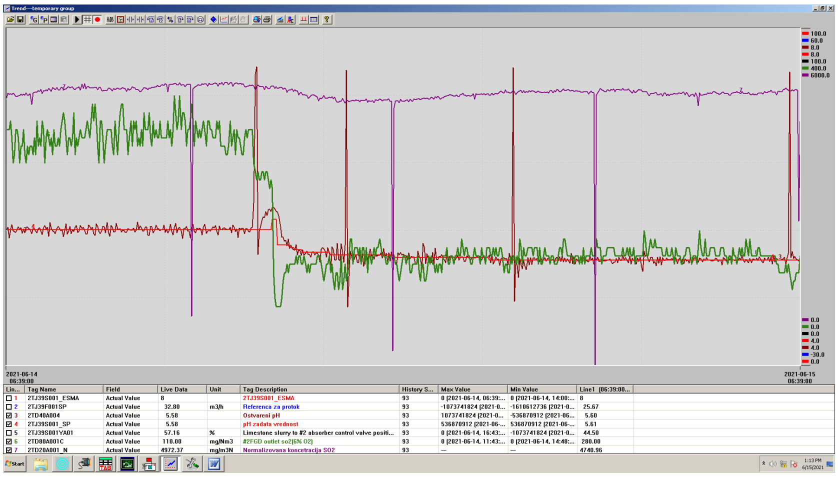


Fig. 8 Control performance of Absorber unit - 24hours time scale: pH reference value – red, pH measured value – brown, SO_2 inlet concentration [mg/Nm^3] – violet, SO_2 outlet concentration [mg/Nm^3] – green

V. Conclusion

The paper presents the application of the cascade control structure of desulphurization in the thermal power plant TE KO Drmno. By applying the proposed control structure, the efficient operation of the entire system was achieved in terms of meeting the regulations of the European Commission, which is shown through the data obtained in the actual operation of the plant. Possible improvements consist of the introduction of feedforward control depending on the boiler load, as well as the measured input values of sulfur oxide.

Acknowledgment

This research was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

References

[1] Uredba "*O graničnim vrednostima emisija zagađujućih materija u vazduh iz postrojenja za sagorevanje*", ("sl. Glasnik rs", br. 6/2016)

[2] Nacionalni plan za smanjenje emisija glavnih zagađujućih materija koje potiču iz starih velikih postrojenja za sagorevanje, "*Službeni glasnik RS*", broj 10 od 6. februara 2020.

[3] "*Pravci optimalnog smanjenja emisija sumpornih oksida iz termoelektrana EPS-a*" – studija, JP EPS, 2006

[4] M. R. Mataušek, G. S. Kvaščev „*A unified step response procedure for autotuning of PI controller and Smith predictor for stable processes*", Journal of Process Control, vol 13, pp. 787-800, 2003.

БИОМЕДИЦИНСКА ТЕХНИКА
/
BIOMEDICAL ENGINEERING

(БТ/BTI)

EMG feedback for improved control of myoelectric hand prostheses

Strahinja Dosen, Pranav Mamidanna, Jack Tchimino, Filip Gasparic, Nikola Jorgovanovic

***Abstract*—Myoelectric prostheses can compensate for the motor functions that are lost due to an amputation. However, none of the commercial prostheses restores somatosensory feedback to their users. A conventional approach to providing the missing sensory information is to read the data from the sensors embedded in the prosthesis and transmit this information back to the subject by stimulating the skin of the residual limb mechanically or electrically. However, we have proposed a substantially different approach to closing the control loop. In this scheme, the tactile feedback does not convey the output of the prosthesis (sensor data) but its command input, namely, the magnitude of the myoelectric signals generated by the user (so-called EMG feedback). In this lecture, we will explain that this method facilitates the natural proprioceptive feedback from the muscles (sense of contraction) and thereby allows predictive control of prosthesis grasping force. We will also present results illustrating that EMG feedback outperforms conventional force feedback in terms of accuracy and robustness. Finally, we will outline the potential for further developments of this approach.**

***Index Terms*—myoelectric prostheses, tactile feedback, electrotactile and vibrotactile stimulation, EMG feedback, grasping force control.**

I. Conventional approach to closing the loop

To fully reconstruct the limb that is lost due to an amputation, a bionic prosthesis should restore both motor and sensory functions. Myoelectric control allows an intuitive connection between the brain of the user and his/her prosthesis; however, commercial devices do not transmit somatosensory feedback. Therefore, the amputees do not "feel" their bionic limbs, which might impair performance during prosthesis use as well as the sense of embodiment.

Modern hand prostheses are equipped with sensors and the feedback can be restored by translating sensor data into stimulation profiles that are delivered to the residual limb using mechanical or electrical stimulation [1]. For instance, the magnitude of the grasping force can be associated with the frequency or intensity of electrical stimulation applied to the surface of the skin. The prosthesis user can then learn to interpret the elicited sensations – faster or stronger stimulation indicates larger forces.

Recognizing the importance of sensory feedback in able-bodied subjects, the expectation was that providing such artificial force feedback to an amputee user would significantly improve the performance. The results in the literature are however contradictory [2]. Some studies indeed show an improvement, while in some, the addition of feedback was not beneficial.

Strahinja Došen, Pranav Mamidanna, and Jack Tchimino are with the Department of Health Science and Technology, Aalborg University, Frederik Bajers Vej 7D 9220, Aalborg, Denmark ({sdosen, prma, itc}@hst.aau.dk).

Filip Gašparić and Nikola Jorgovanovic are with the Faculty of Technical Sciences, University of Novi Sad, 6 Trg Dositeja Obradovića, 21000 Novi Sad, Serbia (e-mail: {filip.gasparic, nikolaj}@uns.ac.rs).

II. EMG feedback: an approach inspired by human motor control

In this lecture, we will first explain that the controversial results in literature can be understood by approaching the topic of artificial sensory feedback from the perspective of human motor control [2]. We will then use the same framework to introduce a different approach to closing the loop, so-called EMG feedback [3]. In this method, the tactile stimulation is not used to convey sensor data (prosthesis output); instead, it transmits the magnitude of the myoelectric signal generated by the user as the command input for the prosthesis. Since the prosthesis responds proportionally to the input signal, EMG feedback enables the subject to control the grasping force predictively.

The usual approach for the implementation of EMG feedback is to divide the myoelectric signal into ranges, where each range is indicated by a simple stimulation pattern, e.g., activation of a specific vibration motor from an array of motors placed around the residual limb [4], [5]. To produce a specific force level, the subject then needs to activate his/her muscles and increase the muscle contraction until he/she feels that the desired motor is on. Then, the subject maintains that contraction level, while the prosthesis closes and produces the desired force. Contrary to the conventional approach, the feedback is in this case available even before the prosthesis makes contact with an object, giving enough time for adjustments.

We will present the results from our recent experiments illustrating this approach and showing that EMG feedback outperforms the conventional method in terms of both accuracy and robustness. We will then present the ideas for the further developments of EMG feedback.

References

[1] S. J. Bensmaia, D. J. Tyler, and S. Micera, "Restoration of sensory information via bionic hands," *Nat. Biomed. Eng.*, Nov. 2020.

[2] J. W. Sensinger and S. Dosen, "A Review of Sensory Feedback in Upper-Limb Prostheses from the Perspective of Human Motor Control," *Front. Neurosci.*, vol. 14, Jun. 2020.

[3] S. Dosen, M. Markovic, K. Somer, B. Graimann, and D. Farina, "EMG Biofeedback for online predictive control of grasping force in a myoelectric prosthesis.," *J. Neuroeng. Rehabil.*, vol. 12, no. 1, p. 55, Jun. 2015.

[4] M. A. Schweisfurth, M. Markovic, S. Dosen, F. Teich, B. Graimann, and D. Farina, "Electrotactile EMG feedback improves the control of prosthesis grasping force," *J. Neural Eng.*, vol. 13, no. 5, p. 056010, Oct. 2016.

[5] J. Tchimino, M. Markovic, J. L. Dideriksen, and S. Dosen, "The effect of calibration parameters on the control of a myoelectric hand prosthesis using EMG feedback," *J. Neural Eng.*, vol. 18, no. 4, p. 046091, Aug. 2021.

Wireless Sensing and Control of Actuation for Machines and Humans

Nenad Jovičić

***Abstract*—To a greater or lesser extent, wireless sensors and control systems have been used for decades in various areas of life and work. First implementations in the industry were intended for reliable transmission of simple remote control commands. Manipulating machines without complicated wiring has encouraged the creation of many ideas for applying wireless technologies in biomedical engineering. However, biomedical systems were often unpredictable, the operation in the environment in which they were applied was not always controllable, and as it turned out that simple technology transfer between industry and human applications was not trivial, sometime even inappropriate.**

Expansion of consumer market and development of computer networks and mobile telephony in the last decades led to the rapid growth of many wireless standards resulting in broadband, narrowband, personal, local, global, and other wireless technologies. The development of technology gives a new impetus to its applications in industry and afterward in life sciences and medicine.

The talk is a retrospective of historical, past, present, and future challenges faced by wireless systems used in industry and biomedicine. Typical industrial and biomedical applications will be analyzed where different parameters like high bandwidth, low latency, or high energy efficiency are essential. Insights on how new wireless standards enable us to apply modern methods of big data processing, cloud computing, and artificial intelligence will be given. The special attention will be put on the bottlenecks in domains of security and reliability.

Finally, the theoretical analysis will be supported by two examples of use of Wi-Fi wireless communication, on one industrial and one biomedical system.

***Index Terms*—wearable system, wireless communication, industry, biomedicine, standards, security, reliability**

Nenad Jovičič is Associate professor of Electronics, School of Electrical Engineering, University of Belgrade., Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: nenad@etf.bg.ac.rs).

A Measure of Spasticity Based on the Exponential Fit of the Knee Joint Torque Estimated from the Goniogram During the Pendulum Test

Antonina Aleksić, Dejan B. Popović, *Member, IEEE*

***Abstract*— Pendulum test is a method to quantify the spasticity. We used the goniogram recorded during the pendulum test to estimate the knee joint torque based on the model which considers spastic reflex activity. We fitted the exponential curve $T_h = ae^{-bt}$ to the estimated knee joint torque to calculate the parameters a and b. We compared the scaled value log a/b with the modified Ashworth score. We used 8 sets of data collected in a clinical study with six complete paraplegic subjects. The comparison shows that the ratio a/b correlates with the MAS scores; thereby, can be used as a measure of spasticity. The advantage of using the ratio a/b is that this score is not rater dependent and that the scores are real numbers compared the MAS scores; thereby, providing better resolution of the level of spasticity.**

***Index Terms*—spinal cord injury; pendulum test; spasticity; modelling**

I. Introduction

Spasticity often follows nervous system lesions at the cortical and spinal levels (e.g., spinal cord injury, multiple sclerosis, cerebral palsy). The spasticity is defined as a motor disorder characterized by a velocity-dependent increase in tonic stretch reflexes (“muscle tone”) with exaggerated tendon jerk, resulting from hyperexcitability of the stretch reflex, as one of the components of the upper motor neuron syndrome [1]. For clinicians, the quantification of the spasticity is of interest to select the most appropriate therapy for the patient. The modified Ashworth Scale (MAS) is the most often method to score the spasticity [2]. The MAS based assessment is characterized by interrater and intrarater reliability [3-5]. Therefore, Wartenberg introduced the pendulum test to eliminate the subjective component when assessing spasticity [6]. The model was expanded by Bajd and Vodovnik [7], Bajd and Bowman [8], and mc later Le Cavorzin et al. [9-11].

Our group recently introduced a score termed Pendulum Test (PT) score [12]. The clinical study proved that the PT score is correlated with MAS [12]. In addition, we used the complex biomechanical model of the pendulum type lower leg movements, and introduced the new measure termed SPASticity Scale (SPAS) [14]. Here we present the correlation of SPAS and MAS on data from 48 pendulum tests and show that the SPAS is more sensitive compared with MAS and show the type of spasticity; thereby, that is useful for clinicians.

.

Antonina Aleksić is with Institute of Technical Sciences of the Serbian Academy of Sciences and Arts, Knez Mihailova 35/IV, Belgrade, Serbia.(e-mail: aleksicantonina@gmail.com)

Dejan B. Popović is with Serbian Academy of Sciences and Arts, Knez Mihailova 35, Belgrade, Serbia and Aalborg University, Aalborg, Denmark (e-mail: dbp@ etf.rs).

II. The Method

A. Subjects

The data were recorded within the clinical study reported in [14]. Six subjects signed the consent approved by the Clinic for rehabilitation “Dr Miroslav Zotović", Belgrade, Serbia. The inclusion criteria were: 1) a complete lesion above the Th12; 2) a stable neurological and medical status; 3) no autonomic dysreflexia; 4) no cognitive disorders; and 5) no medical history of hearing or balance disorders. The recordings were done four times during two sessions. This provided 48 data sets for the analysis.

B. Instrumentation

The pendulum test was recorded using a device that consists two units: one unit mounted at the thigh comprising an amplifier for two electromyography (EMG) signals and an inertial measurement unit - IMU (3D accelerometer and 3D gyroscope), and the second IMU mounted at the shank (Fig. 1, left panel). The sampling frequency of the EMG signal was 1kHz. Both units sent a signal to the host computer wirelessly at 1000 samples per second. More about the device can be found in [15].

C. Data processing

All the data processing was done using MATLAB (Mathworks, Natick, USA). The motion can be modeled with the nonlinear model which includes the passive resistance and the active resistance due to involuntary activity of the stretched paralyzed muscles during the pendulum like movements of the lower leg about the immobile knee. The equation of the motion is the following:

$$J\ddot{\varphi} = -B\dot{\varphi} - K\varphi - \frac{1}{2} mgl \sin\varphi + T_h \quad (1)$$

φ is the knee joint angle measured to the gravity line, J is the moment of inertia for the lower leg, B and K are the damping and stiffness coefficients of the linear model of passive resistance at the knee joint, m is the mass of the shank/foot complex, l the length from the knee to sole. The torque *Th* represents involuntary activities of the paralyzed knee extensors and flexors

$$T_h = T_E(\varphi, \dot{\varphi}) - T_F(\varphi, \dot{\varphi}) \qquad (2)$$

We fitted the recorded goniogram φ *(t)* with the minimum quadratic error method with the model (1) and (2) and calculated the knee joint torque *Th*. The next step was the fit the exponential curve to the calculated torque.

$$T_h = ae^{-bt} \qquad (3)$$

The output from the fitting were the numbers *a* and *b*. We than defined the measure SPAS to limit that output to the range of the scorers of the MAS.

SPAS=4 * log|*a/b*| (4)

More detail can be found in Aleksić [14].

D. Procedure

The subject was sitting on a stable desk with the stable back support, with the hip angle φ≈135°. The thigh was resting on a flat surface while the knee was in front of the edge of the table to allow free rotation of the lower leg about the joint. The pendulum test was performed on each subject 4 times in 2 recording sessions. The examiner extended the knee joint to the full extension ($\varphi \approx \pi/2$), released it, and the lower leg started damped oscillations which stopped in vertical position ($\varphi \approx 0$).

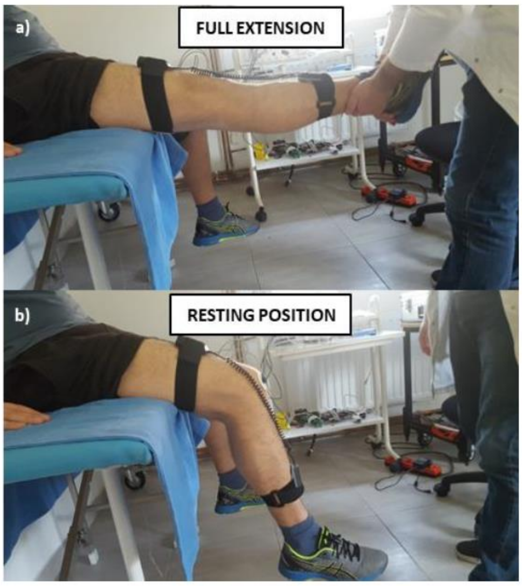


Fig. 1. The setup showing the initial and terminal position of the lower leg during the pendulum test (modified from [15]).

III. Main Results

Figure 1 shows the scaled goniogram of the knee joint by the dashed black line, the calculated moment *Th* shown in red, and the exponential fit in blue. The top two traces show spastic EMG signals that are directly related to the spasticity and the reason of disrupting the normal pendulum like movement in the field of gravity.

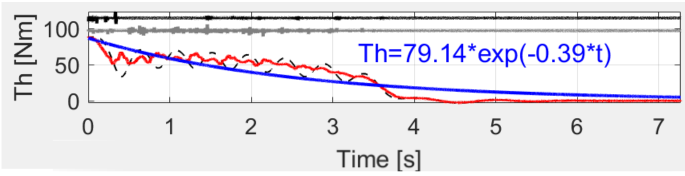


Fig. 2. An example of the estimated torque and the exponential fit for one of the subjects. Red line shows the estimated torque *Th*, the blue line and the equation represent the best exponential fit to the torque *Th*. The dashed black line shows the scaled goniogram of the knee joint $\varphi(t)$ that starts from $\pi/2$ and ends at 0. The black and grey traces at the top are the measured EMG signal (< ±0,5mV) of the knee joint flexors and extensors, respectively.

The MAS and the SPAS values for all the subjects that participated in the study are in Table I.

TABLE I
MAS AND SPAS SCORES FOR THE SIX SUBJECTS ESTIMATED FOUR TIMES IN EACH OF THE TWO RECORDINGS SESSIONS (F - FIRST RECORDING SESSION, S - SECOND RECORDING SESSION)

Nº	1^{st} MAS	1^{st} SPAS	2^{nd} MAS	2^{nd} SPAS	3^{rd} MAS	3^{rd} SPAS	4^{th} MAS	4^{th} SPAS
1f	3	2.43	3	2.24	3	3.24	3	1.30
1s	3	2.97	3	1.52	3	2.89	3	2.31
2f	2	2.62	1	0.67	1	2.02	2	1.63
2s	2	1.83	1	1.90	1	2.01	1	1.55
3f	1	1.19	1	0.78	1	1.22	1	1.50
3s	2	4.80	4	4.85	2	5.13	2	3.66
4f	2	1.75	2	0.78	2	0.74	2	0.98
4s	2	1.16	2	0.08	2	0.41	1	0.32
5f	3	5.68	3	5.52	3	4.76	3	4.88
5s	3	3.78	2	2.21	2	2.94	2	2.35
6f	2	0.63	2	1.00	2	0.34	1	0.08
6s	2	0.31	1	0.17	1	0.17	1	0.17

Data from Table I is presented in Fig. 3. We applied the linear regression and obtained that SPAS = 1.06 MAS -0.12, R^2=0.31 when plotting SPAS vs. MAS.

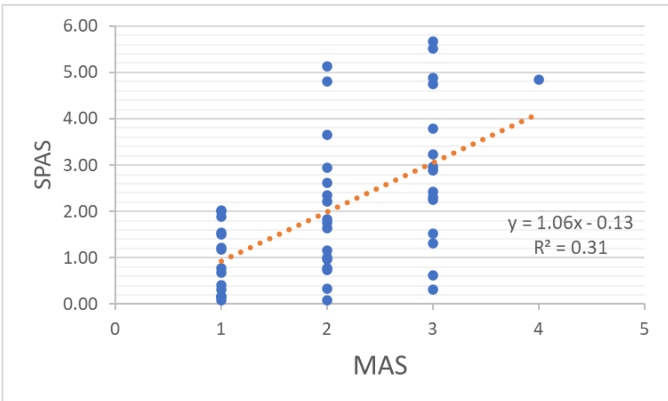


Fig. 3. The SPAS vs. MAS scores and the linear regression for all 48 recording sessions

Figure 4 shows the linear regression of the SPAS vs. MAS without the outliers. The outliers were chosen as the two greatest and lowest SPAS measures for MAS scores 1, 2 and 3. This outlier elimination left for the analysis 36 recordings. The linear regression was SPAS = 0.94 MAS – 0.15, R^2=0.55.

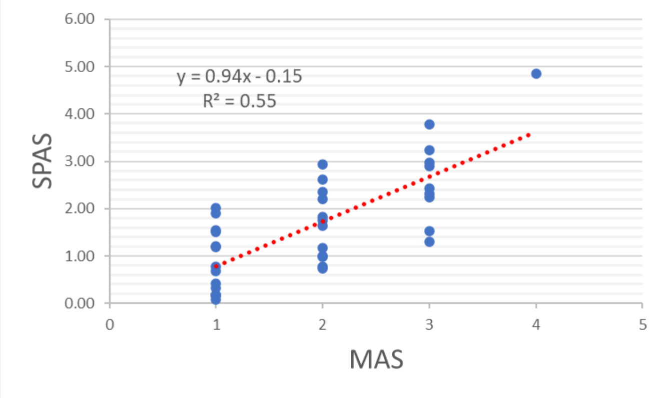


Fig. 4: The linear regression of SPAS vs. MAS of data from 36 recordings.

Figure 5 shows the SPAS and MAS scores from eight recordings for Subject 1 (1f and 1s).

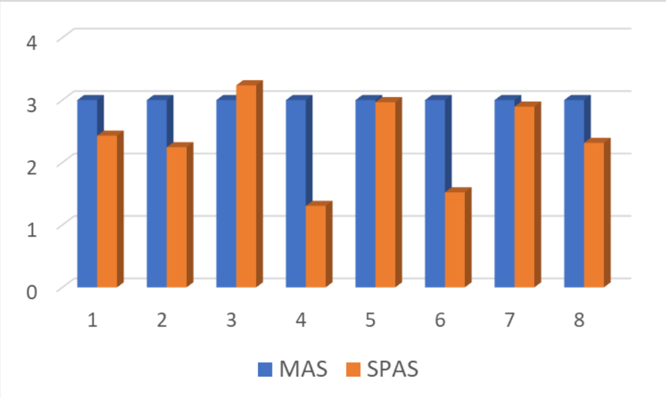


Fig. 5: Bar graph showing the MAS and SPAS scores for Subject 1 (1f and 1s) in eight recordings.

IV. Discussion

Table 1, gives an insight into SPAS and MAS comparison. The graphical presentation of this data (Fig. 3) with the linear fit to has the slope of y=1.06, and R2=0.3. When we exclude 12 outliers, the linear fit equation is y= 0.94x - 0.15, yet with the increased value R^2=0.55. The slope is now smaller compared to the slope with all data, but also close to 1.

In Fig. 5 we show the comparison of the MAS and SPAS for Subject 1. Here, the MAS grade was constant and equal to 3. The SPAS grade varied, being smallest for the fourth recording SPAS=1.3, and highest for the fifth recordings, SPAS=2.97. This clearly suggests how much is the SPAS more sensitive compared with the MAS.

V. Conclusion

We show the relation between the MAS and SPAS scores. Due to the fact that MAS is an integer value and the SPAS is a real number, the sensitivity of the SPAS scale is better compared with the MAS. This increased sensitivity is important for the clinician when deciding which therapy works and following the progress of the recovery of the patient. The SPAS is a measure that automatically determined from the goniogram of the knee joint during the pendulum test.

Acknowledgment

This research was supported by the Ministry of education, science and technology of Republic of Serbia, by the Contract of financing scientific work of Institute of Technical Sciences of the SASA (451-03-9/2021-14/200175) and partly by the Serbian Academy of Sciences and Arts.

We thank PhD Lana Popović Maneski for the support with the instrumentation for the pendulum test and sggestions for the interpretation of data. We thank Radoje Čobeljić, M.D, Ph.D. from the Clinic of rehabilitation "Dr. Miroslav Zotović", Belgrade for the instrumental works with patients and the MAS scoring.

References

[1] [1] J. W. Lance, "Pathophysiology of spasticity and clinical experience with baclofen", In: Lance J. W., Feldman R. G., Young R.R. and Koella W. P. (Eds.), *Spasticity: disordered motor control, Chicago, USA*,pp. 185-203, 1980, Year Book Medical Publishers,

[2] [2] R. W. Bohannon, M. B. Smith, "Interrater reliability of a modified Ashworth scale of muscle spasticity "*Phys ther*, pp. 206-207, vol. 67, no. 2, Feb, 1987.

[3] [3] C. Trompetto, L. Marinelli, L. Mori, E. Pelosin, A. Currà, L. Molfetta, G. Abbruzzese. Pathophysiology of spasticity: implications for neurorehabilitation. *BioMed research international* 2014,

[4] [4] M. Blackburn, P. Van Vliet, S.P. Mockett, "Reliability of measurements obtained with the modified Ashworth scale in the lower extremities of people with stroke",*Physical therapy* vol. 82 no. 1, pp. 25-34, 2002.

[5] [5] A. D. Pandyan, G.R. Johnson, C. I.Price CI, R. H. Curless, M.P. Barnes, H. Rodgers, "A review of the properties and limitations of the Ashworth and modified Ashworth Scales as measures of spasticity",*Clinical rehabilitation*, vol. 13, no. 5, pp. 373-383, 1999.

[6] [6] R. Wartenberg, "Pendulousness of the legs as a diagnostic test." Neurology, vol. 1, pp, 18-24, 1951.

[7] [7] T. Bajd, L.Vodovnik. "Pendulum testing of spasticity." Journal of Biomedical Engineering., vol. 6, no 1. , pp. 9-16, Jan. ,1984

[8] [8] Bajd T, Bowman B. Testing and modelling of spasticity. *Journal of biomedical engineering* vol. 4. no. 2, pp.90-96, 1982.

[9] [9] P. Le Cavorzin, X. Hernot, O. Bartier,H. Allain, G. Carrault,P. Rochcongar, F.Chagneau, " A computed model of the pendulum test of the leg for routine assessment of spasticity in man" , *ITBM-RBM* vol. 22 no. 3, pp. 170-177, 2001.

[10] [10] P. Le Cavorzin, S.A. Poudens, F. Chagneau, G. Carrault, H. Allain, P. Rochcongar P. "A comprehensive model of spastic hypertonia derived from the pendulum test of the leg". *Muscle & nerve*, vol. 24 no.12, pp.1612-1621,2001.

[11] [11] P.Le Cavorzin, G. Carrault, F. Chagneau, P. Rochcongar, ,H. Allain. A computer model of rigidity and related motor dysfunction in Parkinson's disease. *Movement disorders: official journal of the Movement Disorder Society* 2003; vol. 18, no.11, pp. 1257-1265.

[12] [12] L. Popović-Maneski, A. Aleksić, A. Metani, V. Bergeron, R. Čobeljić and D. B. Popović, "Assessment of Spasticity by a Pendulum Test in SCI Patients Who Exercise FES Cycling or Receive Only Conventional Therapy." IEEE Trans Neural Syst Rehabil Eng,, vol. 26, no. 1, pp. 181-187, Jan., 2018.

[13] [13] L. Popović-Maneski, A. Aleksić, R. Čobeljić, T. Bajd, D. B. Popović. "A new method and trumentation for analyzing spasticity." IETI Trans on Ergonomics and Safety, vol. 1, no. 1, pp. 12-27, Jan., 2017.

[14] [14] A. Aleksić, D.B.Popović, "New scale for assessing spasticity based on the pendulum test", Computer Methods in Biomechanics and Biomedical Engineering. in review

[15] [15] M. Miletić, V. Atanasoski, J. Kršić, A. Lazović, and L. Popović-Maneski, "Validation of the new wearable instrument for the pendulum test based on inertial sensors", IcETRAN 2020, 7th International Conference on Electrical, Electronic and Computing Engineering, Belgrade, Serbia, September, 2020.

Multiple measurements by a pendulum test improve the spasticity assessment in SCI subjects

Nikola Babić, Radoje Čobeljić, Slađana Kostić-Smith, Lana Popović-Maneski

***Abstract*— We present the variability of the spasticity scores during three consecutive days using the case series clinical study data with spinal cord injured (SCI) subjects. We assessed the spasticity by the Pendulum Test (PT) and Ashworth Scale (AS) scores. We measured the spasticity on the three consecutive days before and after the period of the treatment. Three subjects with SCI participated in the study. We found large variability from day to day. The PT score had more significant variability compared with the AS. The results suggest that the three consecutive testing by using the pendulum test and PT score on different days provide a better assessment of spasticity being essential in evaluating the treatment protocol.**

***Index Terms*— Spasticity assessment; Pendulum test; PT score; Ashworth Scale; Spinal cord injury.**

I. Introduction

UPPER motor neuron lesions lead to hypertonia resulting in paralysis/paresis of extremities [1]. Spasticity caused by a spinal cord injury (SCI) affects the quality of life of subjects with paralysis. Since neurological recovery is a long process, precise information on spasticity changes could assist in selecting the optimal treatment procedures. A research study that included 110 subjects after SCI shows a significant role in the severity of spasticity of the quality of life when using the Quality of life scores [2].

The validated, most often used methods for assessing the severity of spasticity use the Ashworth scale (AS) and modified Ashworth scale (MAS). AS and MAS score the reaction of manipulating the segment of the body that stretches the muscles responsible for the flexion and extension of the specific joint analyzed. The reaction is assessed physically by the examiner. The fact that the score reflects the examiners' perception of paralyzed muscles' involuntary response makes it inter-rater dependent; therefore, subjective.

The pendulum test (PT) was introduced to get an objective measure of the severity of spasticity. In the PT, the examiner stretches the muscles to bring the body segment into the position where gravity will cause the pendulum-like movement when the examiner lets the body segment free. The PT is convenient for the knee and elbow joints. The PT was used in the severity of spasticity scoring in subjects with spinal cord injuries [3, 4]; cerebral palsy [5, 6]; cerebrovascular insults [3, 7]; and sclerosis multiplex [8].

The PT increases the precision, objectivity, reliability, and validity of the spasticity measures [8]. The essential fact is that spasticity varies; hence, even the quantified PT gives results that change with repetitions and change from one to the next assessment session. We decided to analyze the variability of AS and PT scores by determining those in three consecutive days.

II. Methods and Materials

A. Subjects

The inclusion criteria for entering the study were complete or incomplete SCI above Th12 level (American Spinal Cord Injuries Association (ASIA) A and B), no other trauma or neurological diseases, and the ability to follow the test protocol.

The Ethics Committee of the Clinic for Rehabilitation "Dr. Miroslav Zotović," Belgrade, Serbia, approved a protocol prepared along with the Helsinki declaration. All participants signed the informed consent before entering the study. The basic demography of participants in the study is in Table 1.

TABLE I
BASIC DATA FOR THE SUBJECTS

Subject N^o	ASIA	SCI lesion (motor/sensory)
1	A	Th6/Th6
2	B	C6/Th10
3	A	Th8/Th8

B. Instrumentation

We used an instrument that consists of two units fixed at the thigh and shank. Each unit has the inertial measurement unit (IMU) measuring the acceleration and angular rate of the thigh and shank. The unit at the thigh has a two-channel electromyography (EMG) amplifier system connected to surface electrodes over the knee extensors (Quadriceps m.) and knee flexors (Hamstrings m.) (https://www.3-x-f.com/products.html#121) The reference electrode was placed at the bony part in the vicinity of the knee joint.

Nikola Babić, Ph.D. student, University of Belgrade, Academy of Applied Studies Belgrade, Cara Dušana 254, 11080, Zemun, Serbia (e-mail: splendz@gmail.com).

Radoje Čobeljić, M.D., Ph.D. Clinic for Rehabilitation "Dr Miroslav Zotović", Sokobanjska 13, 11000, Belgrade, Serbia (e-mail: radoje.cobeljic@gmail.com).

Slađana Kostić-Smith M.D., (e-mail: ksladja@hotmail.com).

Lana Popović Maneski, Ph.D. Institute of Technical Sciences of SASA, Knez Mihailova 35.11000, Belgrade, Serbia (e-mail: lanapm13@gmail.com).

C. Protocol

The experienced examiner (one of the authors of this paper) assessed both legs' flexion and extension spasticity by applying the Ashworth scale (AS). After the AS assessment, the physiatrist performed the pendulum test, and the computer calculated the PT score.

All tests were done before the intake of medications for the reduction of spasticity.

The subject's position was between the lying and sitting position, with the shanks hanging over the table edge. Hips angles were set to about 3π/4 degrees in all tests. The pelvis and the thighs were fixed with tape to the table. The clinician extended the knee joint to bring the shank to the horizontal position (full knee extension) and let the shank swing about the knee joint caused only by gravity force. The oscillations stopped in less than 15 seconds in all subjects.

IMUs and EMG signals were simultaneously recorded during the pendulum-like movements until the lower leg stopped swinging. The EMG recordings allow the analysis of the timing and strength of spastic reflexive contractions of the paralyzed knee extensors and flexors. The kinematics recordings were used to calculate the PT score [7]. The pendulum measurements were repeated three times at approximately 15-second intervals. The AS and pendulum tests were repeated on three consecutive days.

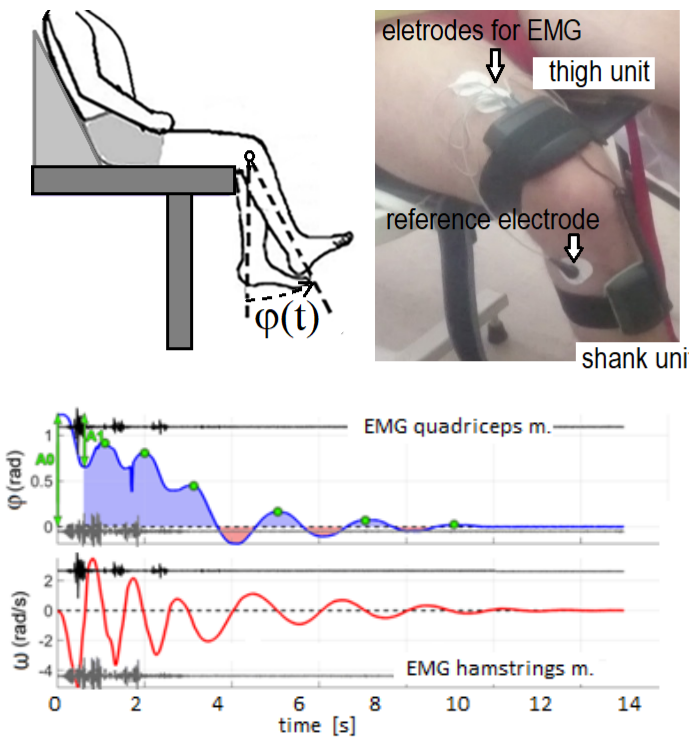


Fig. 1. The setup for the pendulum test with EMG recordings electrodes (upper panel). An example of signals estimated from the recordings: goniogram, tachogram, and the recorded EMG signals (bottom panel).

Fig. 1 shows the knee joint φ(t) and the shank angular rate ω(t) during the pendulum test. The resting position of the shank was vertical (φ = 0). The starting position of the shank in the pendulum test was horizontal (φ = π/2). The blue areas show intervals when the knee joint angle was positive (φ ≥ 0), and the red areas show the intervals when the knee joint was negative (φ < 0). The φ(t) and ω(t) recordings show apparent asymmetry, more precisely, the knee angle is mostly above zero, and there is a significant change in the frequency of oscillations.

EMG signals from the knee extensors (black) and flexors (grey) are shown above and below the goniogram and the tachogram. Fig. 1 shows that the spastic EMG activity stops much before the shank stops oscillating.

The PT score was calculated using the formula defined in Popović Maneski et al. [4]:

$$PT_i = \left|\frac{(\bar{R}_{2n_i}-\hat{R}_{2n_H})}{7*\hat{R}_{2n_H}}\right| + \left|\frac{(\bar{N}_i-\hat{N}_H)}{7*\hat{N}_H}\right| + \left|\frac{(\bar{\varphi}_i-\hat{\varphi}_H)}{7*\hat{\varphi}_H}\right| + \left|\frac{(\bar{\omega}_{max_i}-\hat{\omega}_{max_H})}{7*\hat{\omega}_{max_H}}\right| + \left|\frac{(\bar{\omega}_{min_i}-\hat{\omega}_{min_H})}{7*\hat{\omega}_{min_H}}\right| + \left|\frac{(\bar{f}_i-\hat{f}_H)}{7*\hat{f}_H}\right| + \left|\frac{\left(\left|\overline{\frac{P^+-P^-}{P_{total}}}\right|_i - \left|\widehat{\frac{P^+-P^-}{P_{total}}}\right|_H\right)}{7*100}\right| \quad (1)$$

The index H is used for the values of healthy subjects. "i" denotes a subject N°, ¯ represents a mean value of three trials in the same subject, and ^ represents the mean value for the whole population (i.e., H group population). To normalize PT, each member in the equation is divided by the total number of parameters used to calculate PT (i.e., seven parameters). N is the number of oscillations, $R_{2n} = A_1/1.6A_0$ relaxation index, φ and ω are the knee joint angle and angular rate of the shank, f frequency of oscillations, $P+$ and $P-$ positive and negative area between the knee joint and the neutral value (φ=0), respectively.

III. RESULTS AND DISCUSSIONS

From day to day, we found significant differences in the PT scores. Goniograms with the most significant differences found in the three days of testing are presented (Fig. 2).

Fig. 3 shows the PT and AS scores for three participants in the study for the left and right leg.

The bars indicate that PT scores were different on three subsequent days, while the AS showed more negligible, and in some cases, no variation. We found no rule related to the variability of the PT scores, but this is most likely because the sample of only three subjects is too small. The possible answer will come after a planned randomized clinical study where the statistics are meaningful, and the comparison can be made before and after the treatment.

Figure 4 shows the variability of the results for the measurements on three consecutive days. Although the standard deviation for the left leg in the first and the right in the second day is close to the mean, compared to overall, the data dispersion in the PT scores indicates a higher level of the test sensibility comparing to the AS.

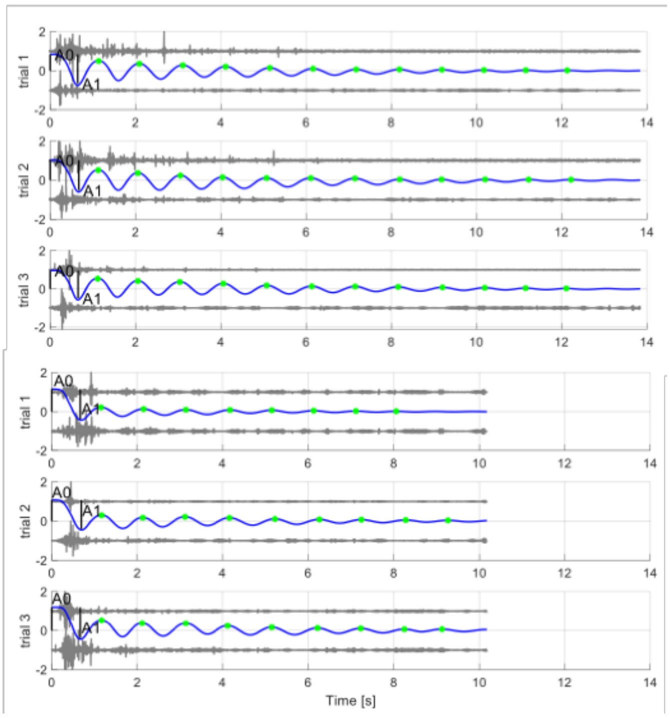


Fig. 2. Goniogram for Subject 1 recorded during the pendulum (right leg). The top three traces are three repetitions during the 1st day, and the bottom three traces are the repetitions on the 2nd day.

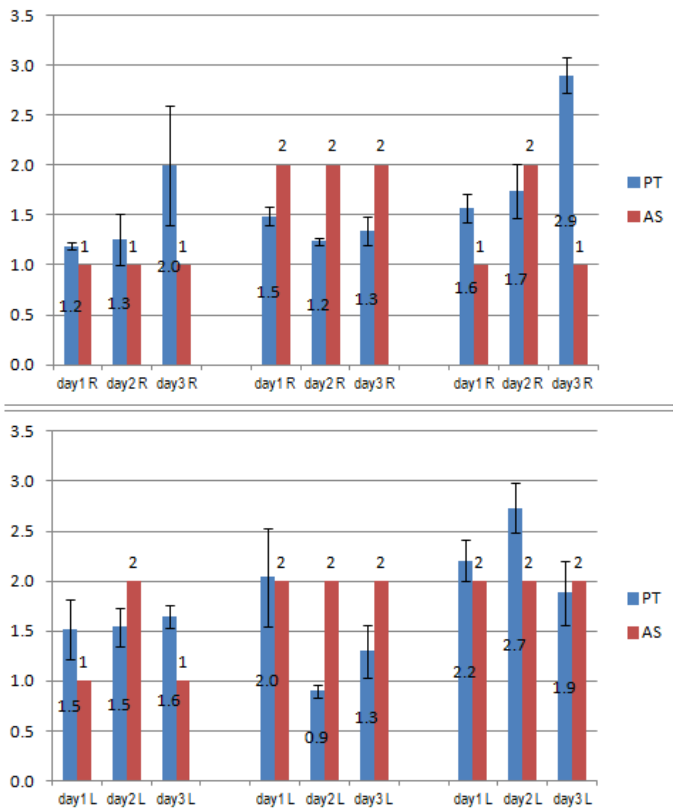


Fig. 3. The PT scores and AS for the right leg (top panel) left and the left leg (bottom panel) for three subjects on three subsequent days.

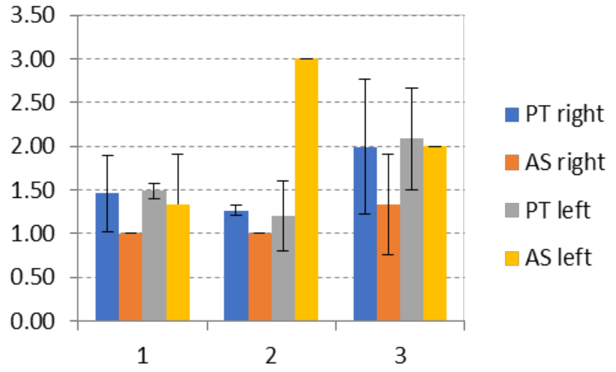


Fig.4. Variability of the PT and AS scores were measured on three consecutive days. The bars show the average score on each day, and the lines are standard deviations.

The size of the study made us present the results by using the means and standard deviations instead of the box-plot presentation.

IV. Conclusion

This case series confirms that precisely detected variability of the spasticity is to be considered in assessing achievements in the rehabilitation progress. Although the variability of the assessment was seen with the Ashworth scale, the precision obtained by the PT score shows that the three consecutive testings on different days are meaningful for the better judgment of spasticity. PT score obtained this way gives precise insight into the spasticity variability, which enables the detection of improvements or deterioration of the subject's condition, thus the nature and extent of effects of a different way of the treatment of spasticity in SCI subjects. The Covid19 pandemic made impossible the inclusion of a more significant number of subjects in the sample and the implementation and evaluation of FES-assisted biking. The findings presented in this paper calls for a more extensive randomized clinical study to be validated.

Acknowledgment

We thank Prof. Dejan B. Popović, from the Serbian Academy of Sciences and Arts (SASA), Belgrade, Serbia, for valuable suggestions and comments related to the final version of the manuscript. We acknowledge the support of physiotherapists Jelena Milovanović and Vladimir Stevanović for performing all the necessary exercises and assistance in measurements. The work on this project was partly supported by the Ministry of Education, Science and Technological Development and contract 451-03-9/2021-14/ 200175 for financing research in ITS-SASA in 2021.

References

[1] C. Trompetto, L. Marinelli, L. Mori, E. Pelosi, A. Currà, L, Molfetta, G. Abbruzzese, "Pathophysiology of spasticity: implications for neurorehabilitation,"Biomed Res Int., 2014.

[2] M. Vural, E.Y. Yalcinkaya, B. Erhan, E.C. Celik, B. Gunduz, A. Bozan, "Assessment of Quality of Life in Relation to Spasticity Severity and

Socio-Demographic and Clinical Factors among Subjects with Spinal Cord Injury, "Journal of Spinal Cord Medicine, vol. 43, no. 2, pp. 193–200, 2020

[3] T, Bajd, L. Vodovnik, "Pendulum testing of spasticity," J Biomed Eng, vol. 6, no. 1, pp. 9-16, Jan 1984.

[4] L. Popović-Maneski, A. Aleksić, A. Metani, V. Bergeron, R. Čobeljić, D.B. Popović. "Assessment of Spasticity by a Pendulum Test in SCI Subjects Who Exercise FES Cycling or Receive Only Conventional Therapy. " IEEE Trans Neural Syst Rehabil Eng." vol. 26, no. 1, pp. 181-187. Jan 2018. DOI: 10.1109/ TNSRE.2017.2771466.

[5] E.G. Fowler, A.I. Nwigwe, T.W. Ho, "Sensitivity of the pendulum test for assessing spasticity in persons with cerebral palsy." Dev Med Child Neurol.; vol. 42, no. 3, pp. 182-189. Mar 2000.

[6] E. Nordmark, G. Anderson. "Wartenberg pendulum test: objective quantification of muscle tone in children with spastic diplegia undergoing selective dorsal rhizotomy." Dev Med Child Neurol. vol. 44. no. 1, pp. 26-33. Jan 2002.

[7] D. B. Popović, T. Bajd. "Pendulum Test: Quantified Assessment of the Type and Level of Spasticity in Persons with Central Nervous System Lesions." Serbian Journal of Electrical Engineering, vol. 15, no. 1), pp. 1–12, Feb 2018.

[8] C.L. Hugos, M.H. Cameron, "Assessment and Measurement of Spasticity in MS: State of the Evidence," Curr Neurol Neurosci Rep, vol 19, no. 79, Aug 2019

[9] P. Decq, P. Filipetti, J.-P. Lefaucheur, "Evaluation of Spasticity in Adults," Operative Techniques in Neurosurgery, Vol 7, no. 3, pp. 100-108, Sep 2004.

Proof of concept platform of an electrotactile Brain Computer Interface

Marija Novičić, Vera Miler-Jerković, Olivera Đorđević, Ljubica Konstantinović and Andrej M. Savić

Abstract—**The aim of this paper is to present the concept and feasibility test of an electrotactile BCI platform consisted of EEG device, electrical stimulation device of nerves/muscles and custom software platform for device control. The developed application comprised GUI for device settings and synchronization of signal acquisition and stimulation control. Experiments for validation of the platform included trancutaneous electrical stimulation at 2 positions on the forearm for inducing somatosensory evoked potentials in the EEG signals in parallel with the tactile attention task performed by the subject. Initial results show that we were able to successfully acquire SEP with our system and that the tactile attention task modified SEP components in a physiologically congruent manner.**

Index Terms—**Brain-computer interface; Event-related potentials; Somatosensory evoked potentials; Electrical stimulation.**

I. Introduction

BRAIN-computer interfaces (BCI) allow the direct link between a person's intentions and technical devices without the need for motor control. BCI devices are promising tools in the domain of assistive technologies for people with motor impairments due to neurodegenerative diseases, spinal cord injuries, stroke or brain trauma [1].

BCI control is based on brain signal measurements such as electroencephalography (EEG). Different EEG based control signals can be utilized for driving BCIs, such as event-related potentials, slow cortical potentials or brain oscillatory activity [2]. Those control modalities should enable the user to activate the function based on different mental strategies.

Event-related potentials (ERP) are commonly used BCI control signal and their features have been used for driving BCI based spellers and menus [2]. ERP based BCIs utilize synchronous BCI control where an ERP is elicited by an external stimuli and voluntary control of the device is based on mental strategy of selective attention towards a single stimulus which results in a mismatch in ERP between attended and unattended conditions which can be detected by BCI.

Since external stimuli for eliciting ERPs can be visual, auditory, and tactile, different types of stimuli have been previously tested for ERP-based BCI control. Most used is the visual modality while the auditory modality is not widely used because of its susceptibility to environmental interferences and relatively low accuracy [1]. The tactile BCIs are not as well-studied probably due to the need for a dedicated stimulation device for ERP eliciting which is more complicated than using a computer screen, light source or a speaker [3].

In this paper, we present an experimental, proof of concept platform of an (electro)tactile BCI based on EEG measurements of somatosensory evoked potentials (SEP) elicited by transcutaneous electrical stimulation (ES) of nerves/muscles.

II. Methods

A. Subjects

One healthy male right-handed volunteer (aged 25) participated in this study. He was without a history of neuromuscular disease and with normal vision and had no previous experience with EEG measurements or proof of concept BCI platform. The study was approved by local ethical committee.

B. Instrumentation and experimental setup

The EEG signals were acquired using the g.USBamp amplifier (g.tec GmbH, Austria) in combination with active (g.GAMMAcap2 connected to g.GAMMAbox, g.tec GmbH) at six recording sites for EEG electrodes arranged according to the 10–20 system: FP1, C3, Cz, C4, CP5 and P3. The reference electrode was placed on the left earlobe and the ground was at the location AFz. FP1 location was used to register eye-movement artifacts. The signals were digitized with a 1200 Hz sampling rate and the amplifier was configured to use Notch embedded filtering with cut-off frequency of 50 Hz.

The electrical stimulation was delivered by an eight-channel electrical stimulator, MOTIMOVE (3F – Fit Fabricando Faber, Serbia). This stimulator is fully programmable and allows the change of stimulation parameters such as stimulus amplitude, pulse width and frequency of stimulation and in the case of our proof of concept BCI platform the parameter settings were made by sending commands from PC via USB. The stimulator has an internal battery power supply which allows mobility and isolation from the main supply.

In order to deliver necessary sensory stimuli, 2 stimulation channels were used. One channel is used for electrical stimulation of dorsal surface (stimulus location D), and the other of volar surface (stimulus location V) of the right forearm. Common indifferent electrode for both channels was round of 2.5 cm diameter and placed on the volar aspect of the right wrist. Two round active electrodes of 1 cm

Marija Novičić – School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11000 Belgrade, Serbia (e-mail: novicic@etf.com).

Vera Miler-Jerković – Innovation Center, School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11000 Belgrade, Serbia (e-mail: vera.miler@etf.rs)

Olivera Đorđević – Faculty of Medicine, University of Belgrade, Dr Subotica 8, 11000 Belgrade, Serbia (e-mail: odordev@eunet.rs).

Ljubica Konstantinović – Faculty of Medicine, University of Belgrade, Dr Subotica 8, 11000 Belgrade, Serbia (e-mail: ljkonstantinovic@yahoo.com).

Andrej M. Savić – School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11000 Belgrade, Serbia (e-mail: andrej_savic@etf.com)

diameter were placed over the extensor carpi radialis muscle (D location) and flexor carpi radialis (V location). The stimuli were single pulses (compensated biphasic with exponential discharge current) of 0.25 ms duration of the active phase. Inter-stimulus interval was set to 750 ms.

C. Experimental protocol

The participant was seated in a chair with a computer screen in front of him at a distance of approximately 1 m. To reduce ocular artifacts, the participant was instructed to fix his gaze at a fixation cross in the middle of the computer screen while right arm was resting on the table in front of the subject (Fig. 1).

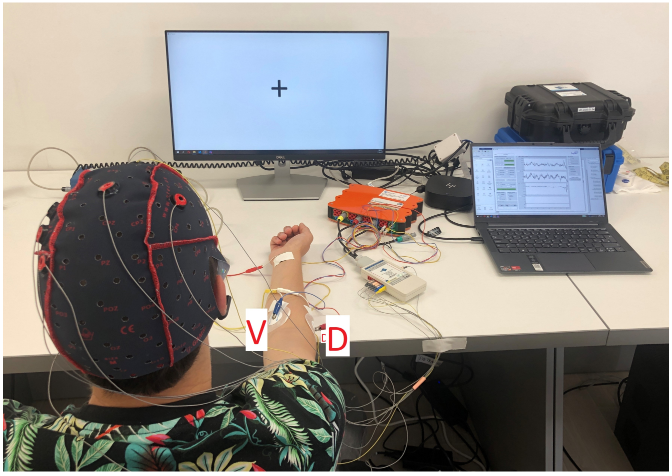


Fig. 1. Experimental setup.

Before the start of the experiment, individual motor threshold was found. Stimulus amplitude was set to 5 mA value and is increased with 1mA step until the contraction was achieved for both D and V locations. The amplitude was then reduced in order to produce the most intense sensation without inducing the contraction. In case of the subject tested those values were 11 mA for D and 10 mA for V.

Experimental protocol was consisted of 6 blocks. Within each block 300 stimuli were delivered in random order to locations D and V while the subject was instructed to attend the stimuli delivered to only one location (D or V) while trying to ignore the stimuli delivered to the other location. The mental strategy in order to maintain the attention was counting the number of stimuli delivered to target location per block. Within blocks 1, 3 and 5 the subject counted the stimuli delivered to location D while in blocks 2, 4 and 6, subject's attention was focused on location V. Therefore, within the duration of the experiment the total number of delivered stimuli was 1800, i.e. 900 per experimental condition (focus on D – F_D, and focus on V – F_V).

D. Software

Graphical user interface (GUI) for stimulation control and data acquisition was developed in MATLAB R2020a (MathWorks Inc., Natick, USA) programming environment (Fig. 2). The GUI consists of four sections.

1) Electrical stimulator control: In this section communication parameters were specified. When communication with stimulator was established, parameters for electrical stimulation were defined. The operator can set value of stimulus amplitude, pulse width and stimulation electrode for both location D and location V. These parameters could be changed during the experiment.

2) Data acquisition: In this section operator needed to set sampling rate and the buffering block size. Sample rate was set to 1200 Hz and buffering block size was set to 60. With these settings, and timer interrupt period, each timer epoch acquires 900 samples.

3) Experiment protocol settings: In this part of the interface operator had to set the parameters that define number of stimuli per block for both channels, number of blocks and duration of pause between blocks. Operator could also choose how the sequence of stimuli was generated. Five options were available: only D is stimulated, only V is stimulated, alternating stimulation between D and V, pseudo random manner with restriction that no more than two consecutive stimuli can be delivered on the same location, and pseudo random manner with restriction that no more than three consecutive stimuli can be delivered on the same location. Finally, operator initialized the session by clicking one of two buttons, depending on which stimulation location the subject has to focus on, so that the dataset was saved and named accordingly to the task (experimental condition).

4) Data visualization: Signals from all channels were shown on the GUI with one of three options. First option showed raw signals. Second one showed signals after filtering with Butterworth 2^{nd} order bandpass filter in a range 1 - 30 Hz. Both first and second option showed signals in time interval of 0.75 s. Third option showed signals of longer duration which was used while preparing the subject for the test and inspecting the EEG signal quality.

E. Data processing

The collected EEG data was bandpass filtered using a 2^{nd} order Butterworth filter in a range 0.1–25 Hz. EEG was segmented to 500 ms epochs (100 ms pre-stimulus baseline and 400 ms post-stimulus interval). Epochs containing artifacts were rejected, where epochs with high absolute amplitude potential shifts (at channels selected for further analysis) and eye-blink/movement artifacts (detected from the Fp1 channel) were selected for rejection. Noise-free epochs were baseline corrected and averaged to form 4 SEP waveforms derived from 2 experimental conditions: 1) focus attention on D while D was stimulated (F_DS_D), 2) focus attention on D while V was stimulated (F_DS_V), 3) focus attention on V while V was stimulated (F_VS_V) and 4) focus attention on V while D was stimulated (F_VS_D). The SEP difference waveforms were calculated by subtracting the SEPs of location D and V delivered within the same condition (task). Namely, difference waveform for FD condition was calculated as F_DS_D-F_DS_V while the difference waveform for F_V condition was calculated as F_VS_D-F_VS_V.

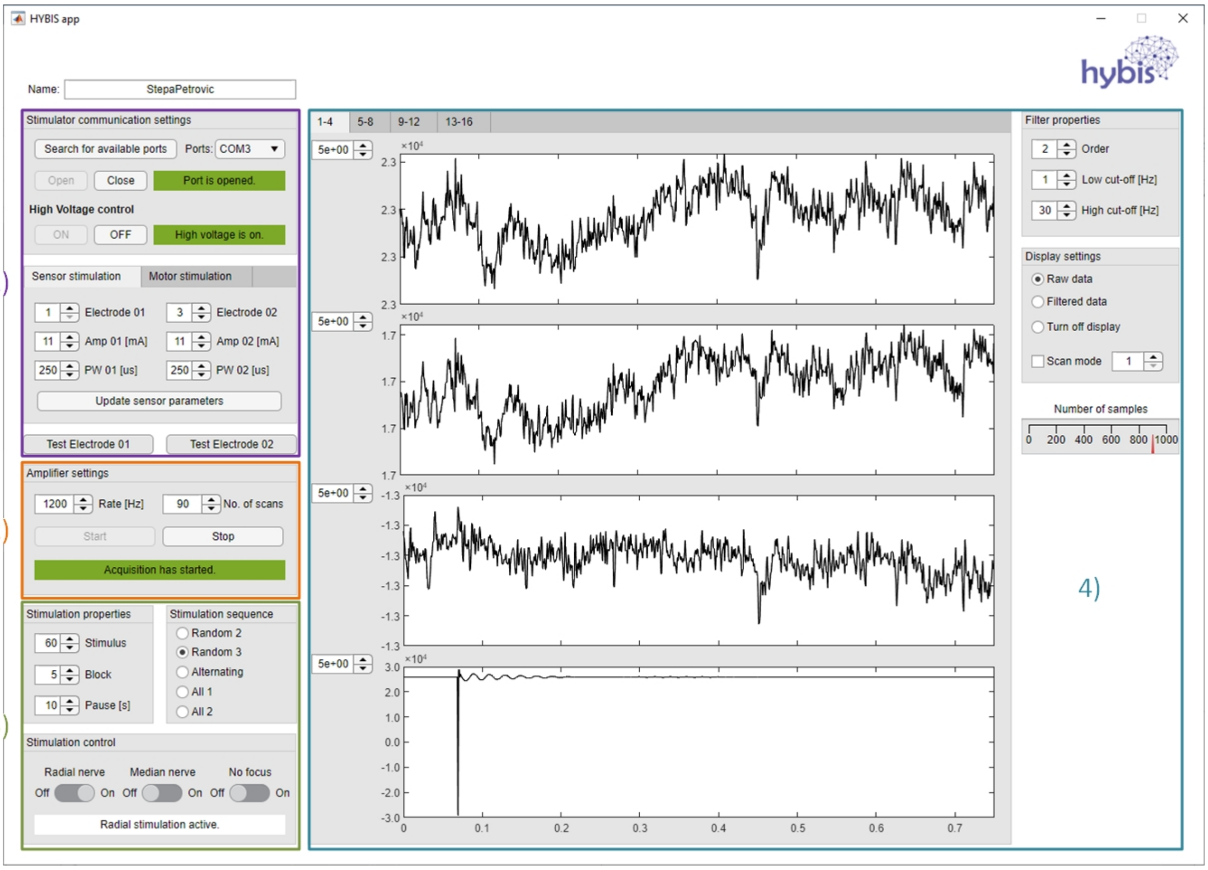


Fig. 2. Main window of the software application GUI during experiment. GUI is divided into four sections: 1) Setting the parameters for communication with electrical stimulator (com port) and stimulation parameters (stimulus amplitude, pulse width and stimulation electrode for both D and V locations), 2) Setting the parameters for data acquisition (sampling rate of acquisition and buffer block size), 3) Setting the parameters for experimental protocol (number of blocks, number of stimuli per block, pause duration and type of sequence), and 4) Data visualisation.

III. RESULTS

SEP waveforms showed that attending the stimuli delivered at one location (D or V) modified the shape of the signal which was reflected in the SEP difference waveform. Representative data for P3 channel is shown in Fig. 3 and 4. Fig. 3 shows the average SEP waveforms when the attention focus was on

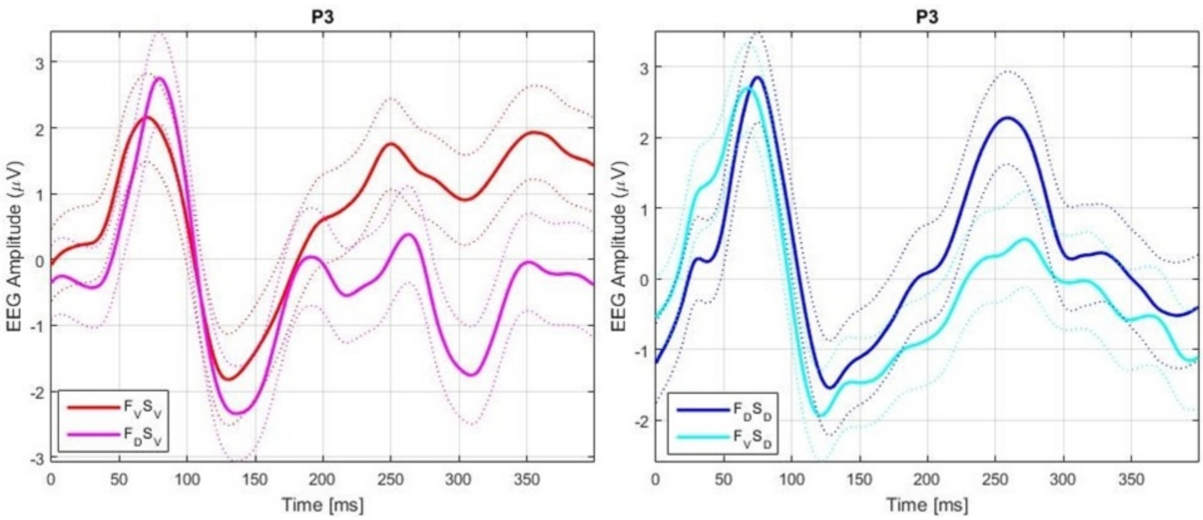


Fig. 3. Left graph shows 2 average SEP waveforms associated with stimulation of dorsal surface of the forearm (SD) while the subject was counting stimuli delivered on dorsal (focus dorsal – F_D, blue line) or volar (F_V , cyan line) surface. Right graph shows 2 average SEP waveforms associated with stimulation of volar surface of the forearm (S_V) while the subject was counting stimuli delivered on volar (focus dorsal – F_V, red line) or volar (F_D, magenda line) surface. Dotted lines present 95% confidence intervals. SEP graphs are presented for EEG channel P3.

the stimulated spot (blue line for D and red for V) or on the other spot (cyan and magenta lines, respectively).

Fig. 4 shows average difference wave between SEP associated with stimulation of dorsal (D) and volar (V) surface of the forearm while the attention focus was on the stimuli delivered to D (blue line) or V (red line).

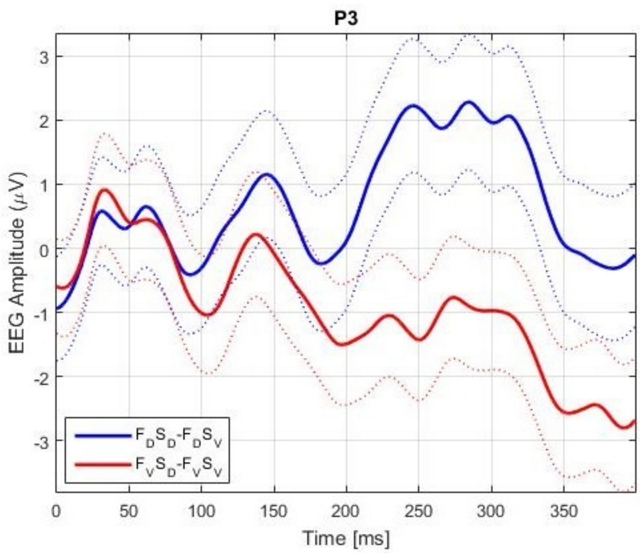


Fig. 4. Blue line represents average difference wave between SEP associated with stimulation of dorsal and volar surface of the forearm while the attention focus was on the stimuli delivered to dorsal surface. Red line represents average difference wave between SEP associated with stimulation of dorsal and volar surface of the forearm while the attention focus was on the stimuli delivered to volar surface. Dotted lines present 95% confidence intervals. SEP graphs are presented for EEG channel P3.

Results indicate significant increase in SEP amplitude of endogenous components, associated with attention focus in both stimulated locations.

IV. Discussion and conclusion

The presented SEP waveforms for both stimulated locations show similar morphology characterized with a first positive component between 50 and 100 ms and negative component between 100 and 150 ms post-stimulus.

Attention focus has resulted in increase of SEP amplitude peaking around 260 ms for SEP associated with D and around 310 ms for SEP associated with V (Fig. 3). The SEP difference waves reveal the window of significance in which the attention focus significantly impacts the processing of somatosensory stimuli in this subject between 210 and 350 ms (Fig. 4) which coincide with P300 ERP component, reflecting the processes involved in stimulus evaluation or categorization [4]. Therefore, our preliminary results validate the SEP recording using our platform and verify that the association of attention focus on SEP amplitude is in a physiologically congruent manner.

This proof of concept platform is a first step towards a novel BCI system for training of somatosensory functions.

Acknowledgment

This work was conducted within project HYBIS funded by the Science Fund Republic of Serbia, under the program for excellent projects of young researchers (PROMIS).

References

[1] C. Chu, J. Luo, X. Tian, X. Han, and S. Guo, “A p300 brain-computer interface paradigm based on electric and vibration simple command tactile stimulation,” *Frontiers in Human Neuroscience*, vol. 15, 2021.

[2] A. Rezeika, M. Benda, P. Stawicki, F. Gembler, A. Saboor, and I. Volosyak, “Brain–computer interface spellers: A review,” *Brain sciences*, vol. 8, no. 4, p. 57, 2018.

[3] A.-M. Brouwer and J. B. Van Erp, “A tactile p300 brain-computer interface,” *Frontiers in neuroscience*, vol. 4, p. 19, 2010.

[4] A. Azizian, A. Freitas, T. Watson, and N. Squires, “Electrophysiological correlates of categorization: P300 amplitude as index of target similarity,” *Biological psychology*, vol. 71, no. 3, pp. 278–288, 2006.

Frequency burst modulation outperforms spatial encoding in multi-level vibrotactile stimulation

Nikolina Maravić, Jelena Bulatović, Filip Gašparić, Strahinja Došen, Nikola Jorgovanović

***Abstract*—Haptic or tactile communication refers to communication through touch. Multichannel vibrotactile stimulation is a commonly used interface to provide tactile feedback. The feedback information is delivered to the subject by modulating stimulation parameters. The present manuscript investigates two approaches for encoding of the feedback information. To this aim, two experiments were performed in 20 healthy able-bodied subjects, whose task was to learn to distinguish eight levels of feedback variable using either burst frequency modulation or spatial locations of vibromotor activation. Vibrotactile feedback was delivered through vibration motors placed on the subject's forearm. The experiments consisted of three phases: a familiarization phase, a reinforced learning phase and a validation phase. The main outcome measure was the success rate in discriminating the levels of the feedback variable. The results have shown that burst frequency modulation (72% success rate) outperformed the spatial coding (64%). Therefore, the frequency encoding is the preferred approach in transmitting multilevel feedback information in vibrotactile feedback systems.**

***Index Terms*—vibrations; stimulation; vibromotors; haptic interface; frequency burst modulation; spatial encoding.**

I. Introduction

The sense of touch is one of the most informative senses, and it is instrumental for daily life activities, haptic exploration and social interaction [1]. There are situations in which a person cannot receive tactile information from the environment, for example, in the case of teleoperation or prosthesis use. In these cases, the missing information can be provided through artificially designed haptic feedback [2]. Generally speaking, the term "haptics" refers to the two types of feedback: feeling of touch on the skin and kinesthetic feedback [3]. Kinesthetic sensations are generated by the sensors located within the muscles and tendons and they allow person to gain a sense of the position of the limbs in space [4].

Nikolina Maravić is with the Faculty of Technical Sciences, University of Novi Sad, 6 Trg Dositeja Obradovića, 21000 Novi Sad, Serbia (e-mail: nikolina.maravic@uns.ac.rs).

Jelena Bulatović is with the Faculty of Technical Sciences, University of Novi Sad, 6 Trg Dositeja Obradovića, 21000 Novi Sad, Serbia (e-mail: jelena996@uns.ac.rs).

Filip Gašparić is with the Faculty of Technical Sciences, University of Novi Sad, 6 Trg Dositeja Obradovića, 21000 Novi Sad, Serbia (e-mail: filip.gasparic@uns.ac.rs).

Strahinja Došen is with the Department of Health Science and Technology, Aalborg University, Frederik Bajers Vej 7D 9220, Aalborg, Denmark (sdosen@hst.aau.dk).

Nikola Jorgovanović is with the Faculty of Technical Sciences, University of Novi Sad, 6 Trg Dositeja Obradovića, 21000 Novi Sad, Serbia (e-mail: nikolaj@uns.ac.rs).

The haptic interface consists of a real-time display of a virtual or remote environment and a manipulator that represents an interface between the human operator and the simulation (VR) and/or remotely controlled system. The user makes movements within a virtual or remote environment by moving the robotic device and those movements are translated to the simulation and/or remote system. Haptic feedback, which is basically force or touch feedback in a human-machine interface, allows computer simulations of various tasks to convey real, tangible sensations to the user, and objects that are typically visually simulated to assume real physical properties, such as weight, hardness and texture. By incorporating haptic feedback into a virtual or remote environment, users have the ability to interact with objects, rather than just see their representation on a monitor [5].

There are two possible ways to reestablish sensory feedback: invasive, by direct stimulation of the physiologically appropriate neural structures in the peripheral or central nervous system, and noninvasive, by stimulating the skin electrically or mechanically [6]. In both cases, the user needs to learn how to associate the delivered stimuli with events and state of the system (e.g., prosthetic hand, gripper of a tele-manipulated robot).

Sensory feedback systems can be divided into three categories: feedback systems based on sensory substitution, feedback systems based on modality-matched stimulus, and somatotopic feedback systems [7].

Sensory substitution is a method that allows information from the environment to reach the user's body through sensory channels that are not intended for that particular stimulus (for example, replacing the sense of touch with the sense of hearing) or through the same sensory channels but when the stimulus arrives in another form (for example, pressure replaced with vibration) [8]. Most feedback systems use this idea, since it is simple to implement. The leading techniques are vibrotactile and electrotactile substitution, which delivers either mechanical vibrations or electric current to the skin to encode informations from the environment [9].

Vibrotactile feedback is one of the most commonly used solutions. In prosthetics, for instance, vibromotor is often used to produce continuous or discrete vibrations when the prosthesis contacts the object [10–12]. The feedback information (e.g., grasping force) can be conveyed by modulating the stimulus frequency [11–13], amplitude [12] or location [13].

In [14], four questions related to vibrotactile feedback for prosthetic hand control were investigated: optimal location for vibromotors, type of signal that activates them, period after which the feeling of irritation decreases after constant stimulation exposure as well as the effect of feedback on grip force control. This study confirmed the improvement in grip force control with the help of vibrotactile feedback.

In the experiment described in [15], 18 healthy subjects operated a virtual object using visual and/or vibrotactile feedback. They received informations via vibration on a finger, hand, neck or foot. All subjects had improved performance when vibrotactile feedback was provided.

The performance of 10 amputees during virtual grasping of objects with feedback on hand aperture and force was investigated in [16]. Their task was to capture the object displayed on the computer monitor with a virtual hand, adjusting the aperture of the hand and the force of the grip using computer mouse. The percentage of correctly applied levels of hand aperture and grip force showed that the use of vibrotactile feedback led to an improvement in hand control compared to the control without feedback.

In [17], two experiments were described: the first referred to the spatial discrimination of stimuli, and the second to the observation of different stimulation intensities. By combining three intensities and three durations of vibrotactile stimulation, nine different stimuli were obtained, which were tested using six vibromotors arranged in four different ways. In the first experiment, circularly placed vibromotors around the upper arm with a proportional distance gave the best results with the accuracy of 75%. Another experiment showed that the perception of vibration intensity was affected by both intensity and duration of vibromotor activation. Seven amputees achieved the accuracy of up to 92% with a circular-proportional vibromotor arrangement.

Despite many studies have used vibrotactile stimulation, no comparison of frequency and spatial coding shemes has been introduced so far. In this study, two vibrotactile coding schemes were presented and compared. Both schemes have the same resolution, they encode 8 levels of vibrations. The difference between two methods is that one implies a variable stimulus burst frequency and other implies variable location of stimulus. The novelty is in comparing frequency and spatial coding schemes which have not been presented so far. The quality of the coding schemes was evaluated using an average success rate achieved by 10 subjects in each experiment in distinguishing levels of vibromotor activation.

II. Methods

The aim of the research described in this paper is to find an adequate way for conveying information using vibrotactile feedback. The idea is to use vibromotors which will be activated according to different spatial and frequency coding schemes so that the user can interpret transmitted feedback information as good as possible. To this aim, we have investigated how well able-bodied subjects could distinguish eight levels of vibrotactile feedback when they are conveyed using different locations versus burst frequency of vibromotor activation.

A. *Experimental environment*

Eight coin type vibration motors (10mm diameter) were installed in the bracelet and placed circumferentially around the subject’s forearm, 2 cm below the elbow (Fig. 1). Vibromotors were marked with numbers 1 – 8. Vibromotor marked with the number "1" was placed in the middle of the lateral forearm, while the others were placed equidistantly, in a clockwise direction. The number of used vibromotors varied depending on the experiment that was performed. In the Experiment 1, four vibromotors were used (marked with numbers 1, 3, 5 and 7), while in the Experiment 2 all eight vibromotors were used. They were connected to the custom made driver board developed at the Faculty of Technical Sciences, University of Novi Sad, which was connected to a PC using a USB cable. The MATLAB software package (version R2018a, MathWorks, USA) was used for creating custom scripts to control the board and collect the data.

Figure 1 shows subject during the experiment. The bracelet with built-in vibromotors was placed around the his/her left forearm, while the subject used his/her right hand to control the mouse when getting acquainted with the levels of vibromotor activation, and then to select the assumed patterns of vibromotor activation.

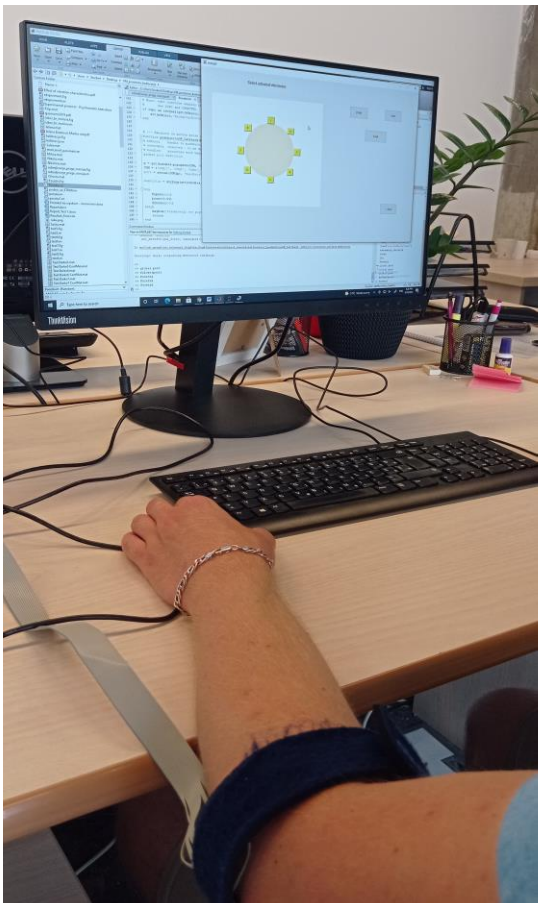


Fig. 1. Subject during the experiment.

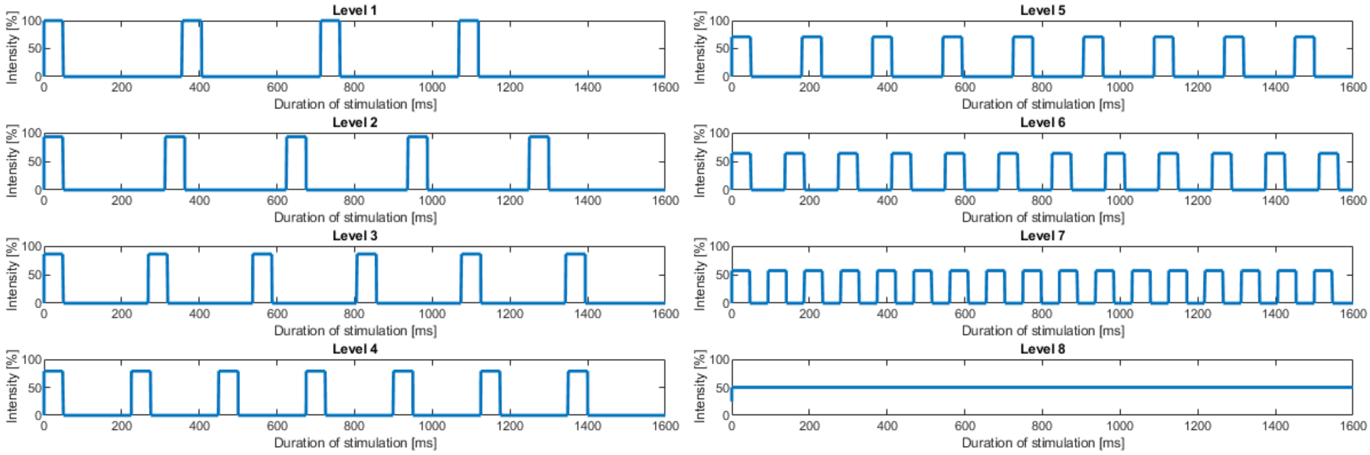


Fig. 2. Control signals for activating vibration motors at eight levels in the Experiment 1.

B. *Stimulation calibration*

Before the beginning of the experiments, it was necessary to determine the parameters of vibrotactile stimulation to create coding schemes that can be clearly perceived by the subjects. For each subject, firstly, the sensation threshold was determined for all vibrators. Each vibromotor was switched on individually, with the vibration intensity gradually increasing from 0 to 100%, in steps of 5%. The intensity at which the subject started feeling vibrations was recorded (sensation threshold).

For the most people the sensation threshold was approx. 30%. Therefore, in the Experiment 1 the intensity of vibrations was obtained by dividing the range between 50% and 100% to eight equidistant values. Remaining stimulation parameters for Experiment 1 were determined after a series of pilot tests. The total duration of stimulation was set to be 1600 ms during which the vibromotors were activated periodically, with an active stimulation ("ON") period of 50 ms. The length of the "OFF" period depended on the level of stimulation and was obtained by dividing the range from 50 to 400 ms by eight equidistant values (Fig. 2). The perceived intensity of vibromotor activation also depends on the length of the stimulation period. By reducing the period of stimulation, i.e. by increasing the frequency with which the stimulus appears, perceived intensity is increasing. However, by adjusting the intensity so that it gradually decreases with increasing burst frequency, the perceived intensity of vibrations can be made approximately equal at each level. When the frequency of stimulus occurrence is the lowest, the stimulation intensity is set to be the highest, i.e. 100% of the maximum value of the simulation, and this pattern was associated to the first level of hypothetical feedback variable. By testing different intensities, it was determined that, when burst frequency is the highest, the activation intensity should be reduced to 50% of the maximum - this corresponded to the eighth level of feedback variable. Figure 2 illustrates the control signals which activate vibromotors. As explained above, the designed patterns produce sensations that modulate in frequency while maintaining approx. the same intensity. Within the pilot tests, the configurations with 1 and 4 equidistantly arranged vibromotors were tested. Subjects reported that the different frequency levels can be distinguished better in the case of using 4 vibromotors. According to that, during the Experiment 1 at a time all four vibration motors were activated simultaneously.

In the Experiment 2, vibrotactile stimulation was delivered through 8 vibration motors equidistantly arranged around the forearm as shown in Figure 3. The vibration intensity was constant and set to 35% to assure that the sensations can be clearly perceived by the subject and yet limit the spread of sensations between adjacent vibrators Only one vibration motor was active at a time. In each trial, the vibromotor was activated for 1000 ms.

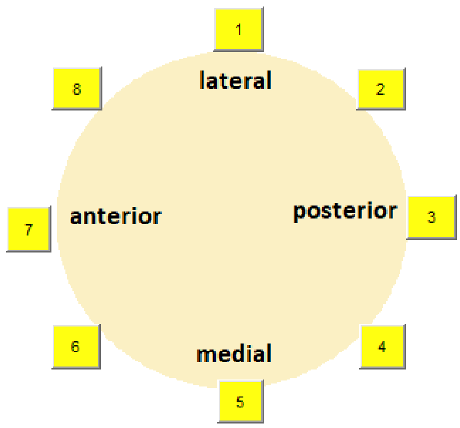


Fig. 3. Locations od vibromotors in the Experiment 2.

C. *Experimental protocol*

The procedure for both experiments was the same and the only difference is that patterns in one case are burst frequency modulations and in the other individual vibrations at different locations. Experimental procedure comprising three phases: the familiarization phase, the reinforced learning phase and the validation phase. Experiments are organized as follows:

1) *Familiarization phase*

During the familiarization phase, the subject was introduced to the experimental environment and to the sensations elicited by different patterns of vibromotor activations. The familiarization phase was finished when the subject was able to distinguish eight patterns of vibromotor activation.

2) Reinforced learning phase

In the learning phase, the vibromotors were activated at different patterns in random order. The subject guessed the value of the activation level after stimulation, and after selecting the answer, the actual value of the level was shown on a computer screen. Each pattern was activated three times in case of frequency coding and ten times in case of spatial coding.

3) Validation phase

The validation phase followed the same protocol as the learning phase, but without feedback information about the correct answer. Validation phase during the Experiment 1 consisted of 25 trials, which means each level appears at least three times (7 levels appears 3 times and 1 level appears 4 times as a task). Experiment 2 consisted of 80 trials, so each location of vibromotor appears 10 times. After 25 (Experiment 1) or 80 (Experiment 2) stimulations, the test phase was completed.

D. Subjects

The subjects in the experiments were professors, assistants and students from the Department of Automation and Computer Science at the Faculty of Technical Sciences, University of Novi Sad. The Experiment 1 was performed in 10 healthy subjects, seven women and three men; aged 29 ± 11 years (mean ± standard deviation), and the Experiment 2 was performed in 10 healthy subjects, six women and four men; aged 31 ± 12 years (mean ± standard deviation). All subjects signed a written consent form to participate in the experiment.

E. Data analysis

During the validation phase, data containing true and predicted values of the location and level of activated vibromotors were collected. The subject had the role of a classifier who had the task to classify each sample, i.e. stimulus into one of the eight possible classes. Here the classes represent the levels of activation of the vibromotor in the first experiment, i.e. locations of activated vibromotors in the second experiment. From the collected data, confusion matrices were formed for each subject. Each row of the matrices represents an actual class, while each column represents a predicted class. All correct predictions are located in the diagonal of the confusion matrices. Based on these matrices, subject’s success rates were calculated as percentage of correct answers among a number of all attempts.

The Kolmogorov-Smirnov test was used to check if data came from a normal distribution. Because of data are not normally distributed, the Mann-Whitney U-test was performed in order to compare two coding approaches. The Mann-Whitney U-test tested the null hypothesis that success rates obtained in two independent experiments are samples from continuous distributions with equal medians. Considering the p-value is 0.2716 we cannot conclude that there is enough evidence to reject the null hypothesis and can conclude that a positive shift in the medians of observed data exists.

III. Results

From the collected data, confusion matrices were formed for each subject. Overall confusion matrices for both experiments were calculated as sum of confusion matrices for all subjects. The performance in recognizing vibration patterns is summarized in Fig. 4 in the form of overall confusion matrices with normalized success rates for all levels.

In Experiment 1, the subjects had the highest success rate in recognizing levels one, two, three and eight, while they were less successful in recognizing other levels. They made the most mistakes in recognizing the fifth level and in distinguishing between levels six and seven. It was easiest for them to recognize level eight, which was expected, considering that in this case the vibromotors were activated continuously for 1600 ms. Also, subjects report that the first, second and third level could be most easily distinguished by simply counting the pulses, since the pulses occur with a frequency low enough to count them.

In Experiment 2, most of the mistakes were confined to adjacent vibromotors. The highest success rate in localizing vibrotactile stimulation was achieved on the lateral side of forearm (vibromotors marked with numbers “1”, “2” and “8” in Figure 3). The lowest success rate in localizing vibrations was achieved on the posterior side of forearm - vibromotor marked with number “3”.

Average success rate for all levels in Experiment 1 is 72%, while in Experiment 2 it is 63%. Therefore, it can be concluded that better results were achieved when feedback variable in vibrotactile systems is stimulus burst frequency than spatial location of the stimulus.

IV. Discussion

A comparison of frequency burst modulation and spatial encoding of vibrotactile stimulation was presented. Twenty healthy subjects participated in two experiments. Effectiveness of each coding scheme was evaluated using success rate achieved in distinguishing eight different levels. Depending on the experiment, stimulus frequency or location was modulated. In the Experiment 1 four vibration motors were activated simultaneously with variable stimulus frequency, while in the Experiment 2 at a time one of eight vibration motor was activated and subject’s task was to localize it. Better results in distinguishing eight different patterns were achieved in case of using frequency coding scheme.

The obtained results make a contribution to the field of haptic interfaces. Vibrotactile stimulation is simple to implement and it allows informations from the environment to reach the user's body through mechanical vibrations. In hand prosthetics, for instance, vibration pattern can be used for identifying the prosthesis contact with the object, or level of grip force.

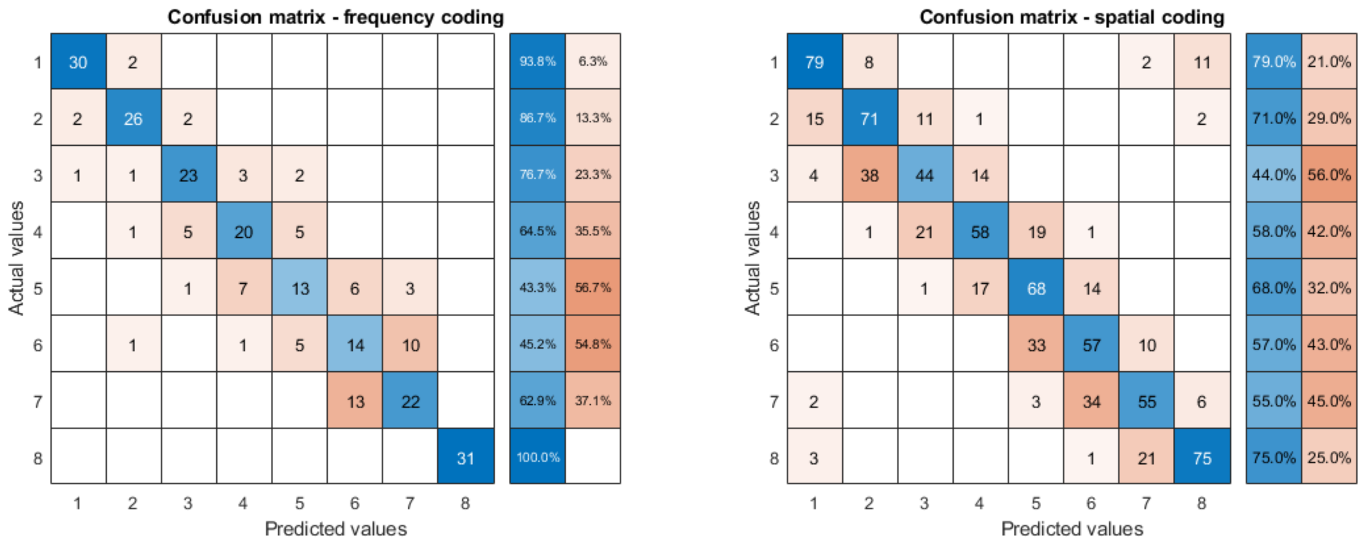


Fig. 4. Overall confusion matrices for both experiments.

V. Conclusion

High performance achieved in described experiments indicate that the frequency and spatial coding schemes of the vibromotor activation are intuitive to use, and they could be applied for providing vibrotactile feedback about 8 levels of feedback variable. However, the frequency burst approach outperformed the spatial encoding. Most of the mistakes made by the subjects involve adjacent levels.

Several female subjects characterized the feeling of vibration on the skin as unpleasant, especially when activating vibromotors at higher frequency levels (6, 7 and 8) during the Experiment 1. The reason for this may be that women's skin is thinner, more moist and covered with less hair, as well as the smaller volume of a woman's forearm compared to a man's, and the intensity of vibration felt in this case is higher. This problem can be solved by adjusting the vibration intensity before starting the experiment for each subject separately.

VI. Acknowledgment

This research was supported by the grants III41007 from the Ministry of Education, Science and Technological development of Serbia.

References

[1] O'malley, M.K. and Gupta, A., 2008. Haptic interfaces. *HCI beyond the GUI: Design for Haptic, Speech, Olfactory, and other nontraditional Interfaces*, pp.25-64.

[2] Brown JD, Paek A, Syed M, O'Malley MK, Shewokis PA, Contreras-Vidal JL, Davis AJ, Gillespie RB: Understanding the role of haptic feedback in a teleoperated/ prosthetic grasp and lift task. *World Haptics Conf* 2013:271–276.

[3] Tzafestas, C.S.: Whole-Hand Kinestethic Feedback and Haptic Perception in Dextrous Virtual Manipulation, *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, Vol. 33, No. 1, 2003.

[4] Proske U, Gandevia SC (2018) Kinesthetic Senses. *Compr Physiol* 8:1157–1183.

[5] H. Culbertson, S. B. Schorr, and A. M. Okamura, "Haptics: The present and future of artificial touch sensation," *Annu. Rev. Control, Robot., Auton. Syst*., vol. 1, pp. 385–409, 2018.

[6] S. J. Bensmaia, D. J. Tyler, and S. Micera, "Restoration of sensory information via bionic hands," *Nat. Biomed. Eng*., pp. 1–3, Nov. 2020.

[7] Schofield, J. S., Evans, K. R., Carey, J. P., and Hebert, J. S. (2014). Applications of sensory feedback in motorized upper extremity prosthesis: a review. *Expert Rev. Med. Devices* 11, 499–511.

[8] Antfolk, C. et al. Sensory feedback in upper limb prosthetics. *Exp. Rev. Med. Dev*. 10, 45–54 (2013).

[9] Kaczmarek, K.A., Webster, J.G., Bach-y-Rita, P. and Tompkins, W.J., 1991. Electrotactile and vibrotactile displays for sensory substitution systems. *IEEE transactions on biomedical engineering*, *38*(1), pp.1-16.

[10] Pylatiuk, C., Kargov, A. and Schulz, S., 2006. Design and evaluation of a low-cost force feedback system for myoelectric prosthetic hands. *JPO: Journal of Prosthetics and Orthotics*, *18*(2), pp.57-61.

[11] Chatterjee, A., Chaubey, P., Martin, J. and Thakor, N., 2008. Testing a prosthetic haptic feedback simulator with an interactive force matching task. *JPO: Journal of Prosthetics and Orthotics*, *20*(2), pp.27-34.

[12] Stepp, C.E. and Matsuoka, Y., 2011. Vibrotactile sensory substitution for object manipulation: amplitude versus pulse train frequency modulation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *20*(1), pp.31-37.

[13] Cholewiak, R.W. and Collins, A.A., 2003. Vibrotactile localization on the arm: effects of place, space and age. *Perc Psychoph* 65:1058- 1077.

[14] Chaubey, P., Rosenbaum-Chou, T., Daly, W. and Boone, D., 2014. Closed-loop vibratory haptic feedback in upper-limb prosthetic users. *JPO: Journal of Prosthetics and Orthotics*, *26*(3), pp.120-127.

[15] Stepp, C.E. and Matsuoka, Y., 2011. Object manipulation improvements due to single session training outweigh the differences among stimulation sites during vibrotactile feedback. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *19*(6), pp.677-685.

[16] Witteveen, H.J., Rietman, H.S. and Veltink, P.H., 2015. Vibrotactile grasping force and hand aperture feedback for myoelectric forearm prosthesis users. *Prosthetics and orthotics international*, *39*(3), pp.204-212.

[17] Guemann, M., Bouvier, S., Halgand, C., Paclet, F., Borrini, L., Ricard, D., Lapeyre, E., Cattaert, D. and de Rugy, A., 2019. Effect of vibration characteristics and vibror arrangement on the tactile perception of the upper arm in healthy subjects and upper limb amputees. *Journal of neuroengineering and rehabilitation*, *16*(1), p.138.

Feasibility Test of Activity Index Summary Metric in Human Hand Activity Recognition

Jelena Medarević, Marija Novičić, Marko Marković

Abstract— **Activity monitoring is a technique for assessing the physical activity that the person undertakes over some time. Activity Index is a metric that summarizes the raw measurements from tri-axial accelerometers, often used for measuring physical activity. Our research compared the Activity Index summary metric for different activity groups and hand usages. We also tested the feasibility of the use of this parameter as a classification feature. Data acquisition was done with previously developed system that includes two smartwatches (one on each wrist) and a smartphone placed in the subject's pocket. Raw data from smartwatch accelerometers was used for the analysis. We calculated the Activity Index for labelled data segments and used ANOVA1 statistical test with Bonferroni correction after data normality was determined by the Lilliefors test (modification of the Kolmogorov-Smirnov test). Significant differences were found between cases of hand usage (left, right, none, both) and between some of the activity groups (walking, sitting standing, grasping, pouring, drinking, opening and closing cupboard, and closing bottle), respectively.**

Index Terms—**Activity Index, accelerometry, smartwatches, ANOVA1, Wilcoxon rank-sum**

Jelena Medarević is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: mj203110m@student.etf.bg.ac.rs).

Marija Novičić is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: novicic@etf.bg.ac.rs).

Marko Marković is with the Applied Rehabilitation Technology Lab, University Medical Center Göttingen, Von-Siebold-Str. 3, D – 37075 Göttingen, Germany (e-mail: marko.markovic@bccn.uni-goettingen.de)

Speech vs. Music Classification Based on EEG Spectral Features Using Artificial Neural Networks

Ivan Vajs, Predrag Jekić, Aleksandra Marjanović, Milica M. Janković

Abstract—**The response mechanisms to different neural stimuli are a challenging task in neuroscience research. The auditory activity (response to music, speech, noise, etc.) can cause various emotional and cognitive responses. The neural responses to speech and music are of particular significance since they are almost constantly present in day-to-day life. We present the classification of the reactions to speech and music based on the spectral EEG features. The mean values of four frequency intervals (corresponding to the theta, alpha, beta, and gamma rhythms) were assessed for seven brain regions. These features were then used as the inputs to the classification based on logistic regression and artificial neural networks; both were used to analyze each subject individually and all available data. Feature selection was also performed, and the classification algorithms were trained using all, a half, and a quarter of the features for comparing their importance and variance for each individual and the entire dataset. The best classification accuracy for a single subject was 85.8%, and an accuracy of 67.1% was achieved for all subjects. This result is promising and calls for the analysis of a larger dataset.**

Index Terms—**EEG; artificial neural networks; logistic regression; classification; feature selection.**

I. Introduction

The analysis of the neural responses to different stimuli is quite widespread in the neuroscience research [1]. One area of interest is the analysis of the relationship between the different types of auditory stimuli and brain activity. More specifically, speech and music are found to be of particular significance, as they are present in all cultures and play an important role in everyday life.

Different modalities can be used to track a person's neural response, with the three most commonly used being electroencephalography EEG [2], magnetoencephalography (MEG) [3] and functional magnetic resonance imaging fMRI [4]. These modalities measure the electrical neural activity, magnetic byproducts of neural activity, the changes related to blood flow, respectively, and are used in a wide range of studies regarding the functional analysis of the brain.

In [5], the authors explored the neural response of 15 subjects when exposed to, and when anticipating audio stimuli. The stimuli were selected to either be neutral, or to induce positive or negative emotions and the response was tracked using MEG. Each stimuli category was preceded with a cue tone of a different frequency so that the subject could know what emotion the following stimuli was meant to induce. It was shown that the brain response was different during the exposure to emotion inducing as opposed to neutral sounds, and that the response of a given stimuli was similar to the response elicited by its corresponding cue tone.

An investigation was carried out in [6] to determine whether neural separability between music and speech response could be detected. There were 47 participants that took part in the experiment and fMRI recordings were made during the exposure to short music excerpts and human vocalizations in a pseudo-random order. The results have shown that there is a specific brain region (a region within the anterior superior temporal gyrus) that responds more strongly to music than voice stimuli.

In [7], a study was conducted with the goal of classifying different musical notes based on the EEG response. Five participants took part in the experiment and the event-related spectral perturbation features were extracted and used as the input to the support vector machine classifier. The results of the study showed a 70% classification accuracy for 12 different classes (notes).

The classification of auditory stimuli (English vowels "a", "i" and "u") was conducted in [8]. Eight subjects took part in the experiment and a recurrent neural network combined with Ben's Spike Algorithm encoding was implemented to classify the EEG signals. The accuracy of 83.2% was obtained when using all 64 available electrodes, and an accuracy of 81.7% when using only 10 of the electrodes.

A classification of speech and music audio recordings was performed in [9]. Although not based on neural response, this paper is interesting because it implemented a novel Spectral Peak Tracking approach applied to the audio recording itself, to

Ivan Vajs is with the School of Electrical Engineering, University of Belgrade, and the Innovation Center, School of Electrical Engineering in Belgrade, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mail: ivan.vajs@ic.etf.bg.ac.rs).

Predrag Jekić is with the School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mail: jp205002p@student.etf.bg.ac.rs).

Aleksandra Marjanović is with the School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mail: amarjanovic@etf.bg.ac.rs).

Milica M. Janković is with the School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mail: piperski@etf.rs).

differentiate between music and speech. Very high classification accuracies (above 99%) achieved using deep learning techniques have shown that the complex structure of music and speech can be differentiated quite well.

Considering that there seems to be neural overlap between the brain response to music and speech [10], effectively distinguishing them using EEG could help separate these responses [11]. This can in turn, aid in the accuracy of the classification of the subjects focus point in the cases of exposure to multiple stimuli, which would be helpful in assistive therapies and the design of hearing loss devices [12].

Although the classification of audio stimuli has been attempted, no study was conducted to differentiate speech and music based on the spectral features of the subject's EEG, using artificial neural networks (ANN).

In this study, the classification of auditory stimuli into two categories (speech and music) was performed using a basic logistic regression model, as well as ANNs. The experiment setup and the EEG processing pipeline, along with the ANN architectures are described in Section II, the results are given in Section III, and the conclusion with directions for future work is given in Section IV.

II. Method

A. Experiment setup

Five healthy participants (Age: 31.4±8.8 years) took part in the experiment. All participants have signed the informed consent. A galvanic skin response (GSR) sensor (Mindfield Biosystems, Gronau, Germany) and an EEG cap (EASYCAP GmbH, Wörthsee, Germany) with 24 electrodes, placed in accordance with the 10–20 system connected to the 24-channel Smarting amplifier (mBrainTrain, Belgrade, Serbia) were used for the recording. Electrode M1 was used as an ECG channel and electrode M2 was excluded from the measurement in order to keep the symmetry of the EEG electrodes. In this study only the EEG signals were taken into consideration, with the sample rate of 256 Hz. The participants were asked to close their eyes and listen to the 30-minute-long audio file containing six sets of recordings trials. Each trial lasted four minutes with a one-minute-long interval of silence beforehand. Three types of audio recordings were played within a trial, each lasting for one minute, separated by 30-second silence intervals. One recording set consists of instrumental music, human speech and bird chirping. A single trial is a random permutation of the three mentioned recording categories. In this study, only the responses to the speech and instrumental music were analyzed. For two of the participants (ID2 and ID3) the measurements from the final third of the experiment were excluded due to the reported discomfort of the participants.

B. EEG processing

Firstly, the recorded EEG signals were filtered using a notch filter to remove the power supply noise at 50 Hz. The EEG corresponding to the music and speech stimuli was cut into data snippets using a time window of two seconds and the time stride of two seconds (i.e., non-overlapping time windows). The data snippet was labeled according to its corresponding stimuli. The feature extraction process was performed on each snippet and consists of the following steps. An estimation of the power spectral density (PSD) was performed electrode-wise, denoted as PSD_e (PSD for electrode e). For every PSD_e, a reference $PSD_{e,ref}$ was extracted from the 10-second interval of silence which precedes the particular stimulus recording. The difference between PSD_e and its respective $PSD_{e,ref}$ (denoted as $PSD_{e,diff}$) was then calculated. At this point, the 22 observed EEG channels (their corresponding $PSD_{e,diff}$) were grouped into 7 categories as follows:

- Frontal left: Fp1, F3, F7, Fz, AFz;
- Frontal right: Fp2, F4, F8, Fz, AFz;
- Central: C3, C4, Cz, CPz;
- Parietal left: P3, P7, Pz, POz;
- Parietal right: P4, P8, Pz, POz;
- Occipital: O1, O2;
- Temporal: T7, T8.

The electrodes were grouped according to their position (frontal, central, parietal, occipital, and temporal), with the frontal and parietal regions being split into two hemispheres, since the number of electrodes in each of the hemispheres was sufficient for them to be observed independently. The $PSD_{e,diff}$ of the electrodes in a single category were averaged, thus creating seven $PSD_{g,diff}$ (PSD for group g, $g \in 1 \div 7$). Finally, the mean spectral power of the following frequency bands (i.e., brainwave activity [13]) was estimated for each $PSD_{g,diff}$:

- [4 Hz, 8 Hz] – theta;
- [8 Hz, 12 Hz] – alpha;
- [12 Hz, 30 Hz] – beta;
- [30 Hz, 80 Hz] – gamma.

This resulted in 4 frequency bands × 7 groups = 28 features for the classification algorithms.

C. Classification algorithms

Multiple models were engineered for the purposes of this study and evaluated using 20-fold cross-validation [14]. In the first part of the study, both a logistic regression (used as a baseline algorithm) and an ANN architecture were designed and trained per participant. The same architectures were used with either 7, 14 or all 28 features as inputs. In the cases of 7 and 14 chosen features, the selection was based on the ANOVA F-value estimated on the training set (Fig. 1.) [15].

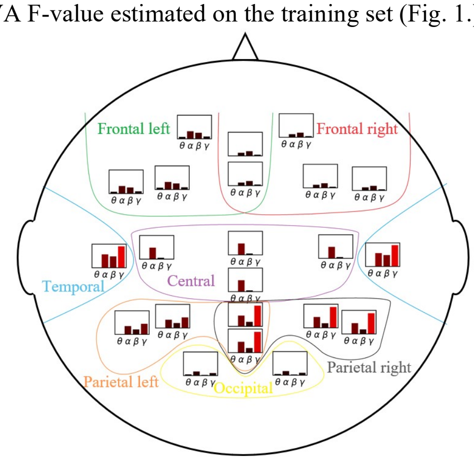


Fig. 1. F-value visualized for each electrode group for a single participant, plotted on the locations of all observed EEG electrodes.

To summarize, there were two different model architectures, three different numbers of input features and five different participants, adding up to 30 different models in total, that were evaluated in this part of the study. In the second part of the study, one logistic regression and three ANN architectures were designed and trained on all collected data, i.e., the data from all participants was placed into a single dataset. These architectures were used for building up different models with either 7, 14 or all 28 features as inputs. Having had four different model architectures and three different numbers of input features, this meant building up 12 different models in total, in the second part of the study.

The shallow ANN used in both the first and second part of the study (ANN1) contained three fully connected (FC) layers (with 5, 4, 2 neurons, respectively). The first hidden layer of ANN1 was a random projection layer, with the purpose of adaptive dimensionality reduction [16]. The other two ANN architectures (ANN2 and ANN3) were evaluated only in the second part of the study. ANN2 contained three FC layers (with 5,10, 2 neurons, respectively). ANN3 consisted of four FC layers (with 5, 15, 10, 2 neurons, respectively). Both ANN2 and ANN3 had a random projection layer as their first hidden layer for the same reason as ANN1. The ANN architectures used for all 28 input features are shown in Fig. 2 (the varying number of input features changes the size of the input layer).

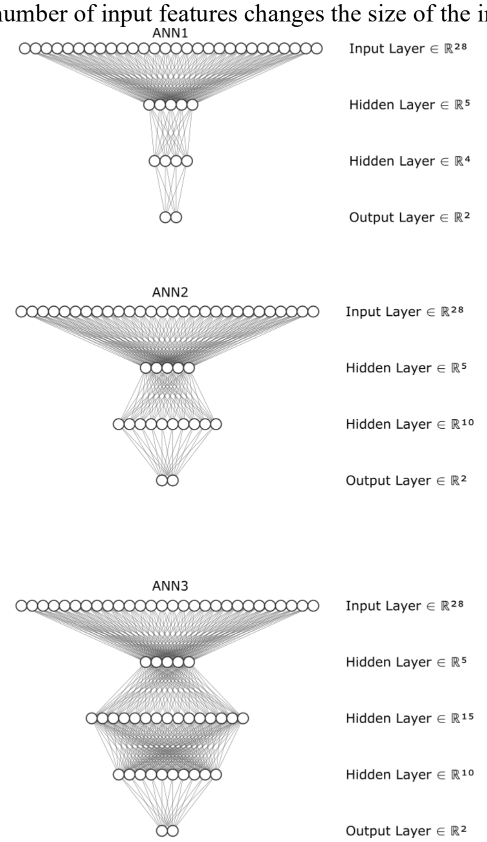


Fig. 2. ANN architectures for all 28 input features.

The relatively low number of neurons and layers was chosen to avoid overfitting considering the dataset size. Adam optimizer was used for the training of the networks with the initial learning rate of 0.00005, batch size of 8, 350 epochs and leaky ReLU activation functions for all hidden layers [17], [18].

III. Results

In Table I, the test classification accuracies are listed per architecture for each subject.

TABLE I
Classification accuracy [%] for each subject.

	Logistic regression			**ANN1**		
	Number of features			**Number of features**		
ID	28	14	7	28	14	7
1	82.9	78.7	78.4	85.8	81.0	79.7
2	74.5	73.3	65.1	75.2	74.2	67.4
3	81.1	75.1	68.4	84.7	76.2	69.5
4	75.6	72.5	69.5	77.6	73.9	70.1
5	74.3	73.4	71.7	74.9	74.3	72.3

The obtained results show that a higher number of input features corresponds to the higher accuracy regardless of the subject and algorithm (which stands in line with the results from [8]). Furthermore, for a given subject and number of input features, ANN1 consistently achieves a higher accuracy compared to the logistic regression. It is important to note that each subject has their specific features that are consistently selected throughout the cross-validation folds and that these features vary between the subjects (Table II). This is expected due to the natural variation of EEG responses between individuals [19].

TABLE II
The selected features for each participant shown for the two folds that exhibit the biggest difference in feature selection [group number followed by brainwave symbol].

ID	**fold**	**14 selected features**	**7 selected features**
1	1	1α, 1β, 2α, 2β, 3α, 4α, 4β, 4γ, 5α, 5β, 5γ, 7α, 7β, 7γ	1α, 3α, 5α, 5γ, 7α, 7β, 7γ
	2	1α, 1β, 2β, 3α, 4α, 4β, 4γ, 5α, 5β, 5γ, 6α, 7α, 7β, 7γ	3α, 4γ, 5α, 5γ, 7α, 7β, 7γ
2	1	2α, 2γ, 3θ, 3α, 3β, 4θ, 5β, 5γ, 6θ, 6α, 6β, 6γ, 7θ, 7γ	2γ, 3θ, 3α, 3β, 5β, 6α, 7γ
	2	1β, 2γ, 3θ, 3α, 3β, 3γ, 5β, 5γ, 6α, 6β, 6γ, 7θ, 7α, 7γ	3θ, 3α, 3β, 4θ, 5β, 6α, 7γ
3	1	1θ, 1β, 2θ, 2α, 2β, 3θ, 3α, 3β, 4β, 4γ, 5α, 5β, 5γ, 6β	3α, 3β, 4β, 4γ, 5β, 5γ, 6β
	2	2θ, 2α, 2β, 3θ, 3α, 3β, 4β, 4γ, 5α, 5β, 5γ, 6θ, 6β, 7α	2θ, 3θ, 3α, 3β, 4γ, 5β, 5γ
4	1	1β, 1γ, 2θ, 2β, 2γ, 3γ, 4α, 4γ, 5α, 5β, 5γ, 6α, 6γ, 7α	1γ, 2γ, 3γ, 4α, 4γ, 5γ, 6α
	2	1β, 1γ, 2θ, 2β, 2γ, 3α, 3γ, 4α, 4γ, 5α, 5γ, 6α, 6γ, 7α	1γ, 2γ, 3γ, 4γ, 5γ, 6α, 6γ
5	1	1θ, 1α, 1γ, 2θ, 2α, 2γ, 3β, 4θ, 4β, 5β, 5γ, 6β, 6γ, 7γ	1θ, 1γ, 2θ, 2γ, 4β, 5γ, 6β
	2	1θ, 1α, 1γ, 2θ, 2α, 2γ, 4θ, 4β, 4γ, 5β, 5γ, 6β, 6γ, 7γ	1θ, 1γ, 4β, 5β, 5γ, 6β, 6γ

In Table III, the test classification accuracies are listed for each model deployed on the set containing the data from all the subjects.

TABLE III
CLASSIFICATION ACCURACY [%] FOR ALL SUBJECTS.

	Number of features		
Architecture	28	14	7
Logistic regression	61.4	59.1	58.3
ANN1	64.8	63.3	62.1
ANN2	65.6	64.2	63.0
ANN3	67.1	64.8	63.9

The overall accuracies shown in Table III are lower than the accuracies obtained when the individual subjects were considered. With respect to the diversity of individual EEG responses and the number of participants it was more difficult for the algorithms to pick up on the complex input-output dependencies.

IV. CONCLUSION

In this paper, the classification of audio stimuli (speech and music) based on spectral EEG features was performed. Firstly, the classification was performed per subject, using the logistic regression and ANN. ANN has shown a slight but consistent improvement (ANN accuracy ranging from 67.4% to 85.8%) over the baseline logistic regression which is to be expected considering the dataset size. Furthermore, a larger number of input features implies a small but consistent increase in accuracy. The deployed models achieved an accuracy above 65% on the test set even when 7 features were selected from the observed dataset. This implies that although a higher number of input features does improve the overall accuracy, certain features do carry more useful information than others. On the other hand, having all collected data in one dataset, resulted in having the maximum accuracy of 67.1%. This is due to the difficulty of achieving higher accuracies when there is an undeniable diversity in the dataset compared to the number of instances and a varying importance of a single feature between subjects.

The directions for the future work include expanding the dataset with significantly more subjects, thus enabling the development of more complex algorithms, alongside the implementation of other EEG processing and feature selection methods. By expanding the database and expanding the EEG feature set, a higher distinction accuracy between speech and music response could be expected. This would open up a possibility to estimate the focus of a given subject when exposed to these stimuli simultaneously, which is often the case in day-to-day life. Further research will also include emotional aspects based on the consideration of heart rate variability (HRV) parameters and the GSR.

ACKNOWLEDGMENT

This research was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia. The authors would also like to thank all the participants that took part in the experiment and dr Andrej Savić for his useful advice regarding the processing of the EEG signals.

REFERENCES

[1] A. F. Meyer, R. S. Williamson, J. F. Linden, and M. Sahani, “Models of neuronal stimulus-response functions: Elaboration, estimation, and evaluation,” *Frontiers in Systems Neuroscience*, vol. 10. Frontiers Media S.A., p. 109, 12-Jan-2017.

[2] C. D. Binnie and P. F. Prior, “Electroencephalography,” *Journal of Neurology, Neurosurgery and Psychiatry*, vol. 57, no. 11. BMJ Publishing Group, pp. 1308–1319, 1994.

[3] S. P. Singh, “Magnetoencephalography: Basic principles,” *Ann. Indian Acad. Neurol.*, vol. 17, no. SUPPL. 1, p. S107, 2014.

[4] G. H. Glover, “Overview of functional magnetic resonance imaging,” *Neurosurgery Clinics of North America*, vol. 22, no. 2. NIH Public Access, pp. 133–139, Apr-2011.

[5] K. Yokosawa, S. Pamilo, L. Hirvenkari, R. Hari, and E. Pihko, “Activation of auditory cortex by anticipating and hearing emotional sounds: An MEG study,” *PLoS One*, vol. 8, no. 11, p. 80284, Nov. 2013.

[6] J. L. Armony, W. Aubé, A. Angulo-Perkins, I. Peretz, and L. Concha, “The specificity of neural responses to music and their relation to voice processing: An fMRI-adaptation study,” *Neurosci. Lett.*, vol. 593, pp. 35–39, Apr. 2015.

[7] K. Tsekoura and A. Foka, “Classification of EEG signals produced by musical notes as stimuli,” *Expert Syst. Appl.*, vol. 159, p. 113507, Nov. 2020.

[8] M.-A. Moinnereau, T. Brienne, S. Brodeur, J. Rouat, K. Whittingstall, and E. Plourde, “Classification of auditory stimuli from EEG signals with a regulated recurrent neural network reservoir,” Apr. 2018.

[9] M. Bhattacharjee, S. R. M. Prasanna, and P. Guha, “Speech/Music Classification Using Features From Spectral Peaks,” *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 28, pp. 1549–1559, 2020.

[10] I. Peretz, D. Vuvan, M.-É. Lagrois, and J. L. Armony, “Neural overlap in processing music and speech,” *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 370, no. 1664, p. 20140090, 2015.

[11] N. J. Zuk, E. S. Teoh, and E. C. Lalor, “EEG-based classification of natural sounds reveals specialized responses to speech and music,” *Neuroimage*, vol. 210, p. 116558, Apr. 2020.

[12] M. Ogg, D. Moraczewski, S. E. Kuchinsky, and L. R. Slevc, “Separable neural representations of sound sources: Speaker identity and musical timbre,” *Neuroimage*, vol. 191, pp. 116–126, May 2019.

[13] C. S. Nayak and A. C. Anilkumar, *Eeg normal waveforms*. StatPearls Publishing, 2020.

[14] M. Stone, “Cross-Validatory Choice and Assessment of Statistical Predictions,” *J. R. Stat. Soc. Ser. B*, vol. 36, no. 2, pp. 111–147, Jun. 1974.

[15] L. Sthle and S. Wold, “Analysis of variance (ANOVA),” *Chemometrics and Intelligent Laboratory Systems*, vol. 6, no. 4. Elsevier, pp. 259–272, 01-Nov-1989.

[16] P. Iwo, Wojcik, “Random Projectionin Deep Neural Networks,” University of Science and Technology in Kraków, 2018.

[17] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv Prepr. arXiv1412.6980*, 2014.

[18] B. Xu, N. Wang, T. Chen, and M. Li, “Empirical evaluation of rectified activations in convolutional network,” *arXiv Prepr. arXiv1505.00853*, 2015.

[19] J. Intriligator and J. Polich, “On the relationship between EEG and ERP variability,” *Int. J. Psychophysiol.*, vol. 20, no. 1, pp. 59–74, Jun. 1995.

How TV commercials affect attention and memory?

Brana Kostić, Vanja Ković, Vera Miler Jerković and Milica Janković, *Member, IEEE*

***Abstract*— Neuromarketing is an emerging multidisciplinary field that involves neuroscience methodology to estimate the reaction of consumers to marketing activities and the way they affect their decisions. The most used neurophysiological technique for neuromarketing studies is electroencephalography (EEG). We present a pilot EEG signals-based study on four participants. We investigated the effect of selected seven commercials on memorization and attention. Statistical analysis of extracted attention and memorization indices has shown high inter-subject variability. It has also demonstrated a statistically significant difference (p<0.05) between participant reactions on commercials on the individual level. Novel metric based on normalized total score of attention index, memorization index and self-assessment was proposed and demonstrated through the comparison of commercials.**

***Index Terms* — neuromarketing, EEG, attention, memory.**

I. Introduction

NEUROMARKETING, a new field of marketing research, has greatly benefited from applying neurophysiological methodology in investigating conscious and unconscious consumer behavior, during the past decade [1, 2]. At the same time, it is overcoming challenges that the traditional metrics used to offer and adding value to the traditional marketing research [3]. Some of the neurophysiological technologies that are used in neuromarketing are: functional magnetic resonance imaging (fMRI, for measuring changes in blood oxygenation and flow that occur as a response to brain activity) [4], electroencephalography (EEG, for measuring electrical activity in the cerebral cortex) [1, 5], eye tracking [6] or various sensors for measuring changes in human physiological condition (e.g., heart rate, respiratory rate, skin conductivity) [7,8]. EEG technology is affordable and with an excellent temporal resolution, so it is the most used neurophysiological technique for marketing studies [1].

The growing literature in this field has unequivocally shown that EEG can be successfully used for research in the field of advertising and neuromarketing. For example, Dimpfel et al. recorded EEG and eye-tracking signals of participants who observed 5 advertisements about banks and concluded that the advertisements affected the observer in a way that the distinct brain regions were activated depending on the type of the emotion evoked whilst the person observed the advertisement [9]. In another EEG study, with 34 participants who observed 5 advertisements, Balconi et al. [10] found a strong correlation between what customers liked and the activation of the dorsolateral prefrontal cortex (DLPFC). In participants who watched commercials, Vecchiato et al. found an asymmetric increase in theta and alpha activity related to the observation of pleasant and unpleasant advertisements, in the left and right hemisphere, respectively [11,12]. More precisely, they found de-synchronization of the left alpha frontal activity, as well as the higher theta activation on the left frontal and pre-frontal areas, if the respondents liked the advertisement. Astolfi et. al. [13] assessed cortical activity in healthy subjects whilst watching TV commercials inserted within a film, using a high-resolution EEG technique, and found highly significant brain activity during the observation of TV commercials, which was mainly concentrated in the frontoparietal parts of the cortex (approximately grouped around Broadman fields 8, 9 and 7).

The goal of this research was to, based on the EEG measurements [14], assess attention and memory indices for each of the commercials presented to the participants in order to rate them according to these markers. We expected that some commercials would score high on the index of memory, some on the index of attention, some on neither of the two indices, and some on both. That way, the practical value of this research would be to point out those advertisements which are better suited for attracting attention vs. those who are memorized better. The second goal was to assess individual differences across participants.

In Section II we have presented the performed methodology: experiment description and EEG data analysis. Furthermore, we have illustrated group and individual statistical results of TV commercials impact on attention and memory, in Section III. Finally, we have given conclusion and plans for the further work in Section IV.

II. The Method

A. Experiment description

Four healthy participants (1 female and 3 males, age: 27.7±7.5 years) participated in the experiment. All participants have signed the informed consent.

Brana Kostić is with the University of Belgrade - School of Electrical Engineering, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mail: branakostic10@gmail.com).

Vanja Ković is with the University of Belgrade, Faculty of philosophy, Laboratory for Neurocognition and Applied Cognition, Čika Ljubina 18-20, 11000 Belgrade, Serbia (e-mail: vanja.kovic@f.bg.ac.rs).

Vera Miler Jerković is with the Innovation Center, School of Electrical Engineering in Belgrade, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mail: vera.miler@etf.rs).

Milica Janković is with the University of Belgrade - School of Electrical Engineering, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mail: piperski@etf.rs).

During the experiment, participants were asked to comfortably sit in front of the computer screen and watch the prepared video. The video consisted of 7 commercials with a 30 s pause (black screen) between them and at the beginning, as well. Commercials were selected from the list of top 2012 commercials in Serbia [15]. We have considered that the advertisements have been shown on TV long ago and that participants do not remember them in detail, so the initial memory effect was negligible. All commercials had approximately same duration and a similar structure, consisting of three parts: 1) story introduction, 2) action and 3) product logo display. The complete video timeline is presented in Table 1.

EEG data acquisition was performed using a 24-channel Smarting amplifier (mBrainTrain, Belgrade, Serbia) connected to Greentek EEG cap (Wuhan Greentek Pty. Ltd, China). Twenty-two monopolar EEG channels were recorded (10/20 locations: Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T7, T8, P7, P8, Fz, Cz, Pz, AFz, CPz, POz). The ground was located at FPz and FCz, and it was used as the reference site. The sampling rate was set to 500Hz. The EEG system sent EEG data via Bluetooth to the computer.

At the end of the experiment, participants were asked to subjectively evaluate each advertisement (scale 1 to 5, 5 was the highest mark).

TABLE I
THE VIDEO TIMELINE

No.	Advertisement's name	Duration
	Pause	30 s
1.	"Forever", Carnex	33 s
	Pause	30 s
2.	"Meeting with happiness", State Lottery of Serbia	30 s
	Pause	30 s
3.	"Infostud jobs"	20 s
	Pause	30 s
4.	"LAV beer"	30 s
	Pause	30 s
5.	"Schweeps"	35 s
	Pause	30 s
6.	"JAFFA cakes"	30 s
	Pause	30 s
7.	"Sportingbet"	28 s

B. *EEG data analysis*

EEG data analysis consisted of 1) EEG preprocessing, 2) extraction of attention and memory indices and 3) statistical data analysis.

EEG preprocessing included FIR filtering in range 2-30 Hz, removing eye-blink artifacts using *Individual component analysis* (*ICA*) method and bandpass extraction of alpha activity (8-12 Hz) in the right frontal (Fp2, F4, F8) and the left frontal lobe (Fp1, F3, F7) and theta activity (4-8 Hz) in the left frontal lobe (Fp1, F3, F7) by a fifth-order Butterworth filter. The flowchart of the EEG preprocessing is presented in Fig. 1.

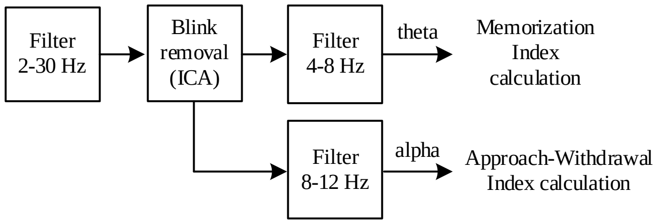


Fig. 1. Flowchart of the EEG preprocessing

Approach-Withdrawal index [7, 8, 16] for attention (*AW*) was calculated on preprocessed alpha activity in the right and left frontal lobes, on time windows of 3 seconds as in

$$AW = P_{\alpha\, right\, frontal} - P_{\alpha\, left\, frontal}$$

$$AW = \frac{1}{N_R}\sum_{i\in R} EEG^2_{r\,\alpha_i}(t) - \frac{1}{N_L}\sum_{i\in L} EEG^2_{l\,\alpha_i}(t) \quad (1)$$

where $P_{\alpha\, right\, frontal}$ and $P_{\alpha\, left\, frontal}$ are the averaged powers of alpha activity in the right and left frontal lobes, respectively; $EEG_{r\,\alpha_i}$ and $EEG_{l\,\alpha_i}$ are the i-th preprocessed alpha band EEG channels in the right (R) and left (L) frontal lobes, respectively; N_R and N_L are numbers of $EEG_{r\,\alpha_i}$ and $EEG_{l\,\alpha_i}$ channels (N_R=N_L=3).

Memorization index (*MI*) [17] was calculated on preprocessed theta activity in the left frontal lobe, on time windows of 3 seconds as in

$$MI = P_{\theta\, left\, frontal} = \frac{1}{N_L}\sum_{i\in L} EEG^2_{l\,\theta_i}(t) \quad (2)$$

where $P_{\theta\, left\, frontal}$ is the averaged power of theta activity in the left frontal lobe. $EEG_{l\,\theta_i}$ are the i-th preprocessed theta band EEG channels, in the left (L) frontal lobe and N_L is the number of $EEG_{l\,\theta_i}$ channels (N_L=3).

EEG signal preprocessing and indices extraction was done in Matlab R2014b (Mathworks, USA). EEGLAB Toolbox was used for filtering and for *ICA* application [18].

The ANOVA analysis was performed on extracted *AW* and *MI* indices to investigate the statistical differences (the significance values of $p<0.05$) in the reactions of the participants on different commercials. We have analyzed changes in *AW* and *MI* indices on the group level (all participants for each advertisement) and on the individual (participant) level for each advertisement. Statistical analysis was done in RStudio [19]. A novel metric defined as a total score for *AW* index, *MI* index and self-assessment was also calculated (as a sum of participant mean values for each advertisement) and normalized to the range [0,1].

III. RESULTS

Box plots for *AW* and *MI* indices on the participant group level for each commercial is presented in Fig. 2. ANOVA analysis has shown statistically significant difference ($p<0.05$)

between commercials only for *AW* index, Fig. 2A. There was no significant difference between commercials on the participant group level in terms of memory, Fig. 2B.

Individual differences between participants concerning their reaction for each commercial were investigated as well. Fig. 3 shows box plots for *AW* (Fig. 3A) and *MI* (Fig. 3A) indices on the participant level for each advertisement. ANOVA analysis has shown statistically significant difference ($p<0.05$) between commercials for both indices, *AW* and *MI*.

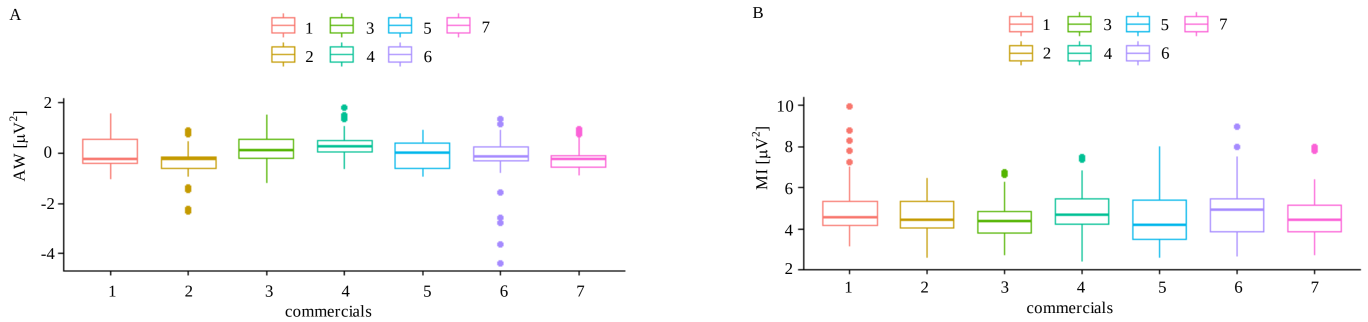


Fig. 2. A) Approach-Withdrawal index (*AW*) and B) Memorization index (*MI*) for each commercial on the participant group level (commercial order 1-7 is shown in Table 1).

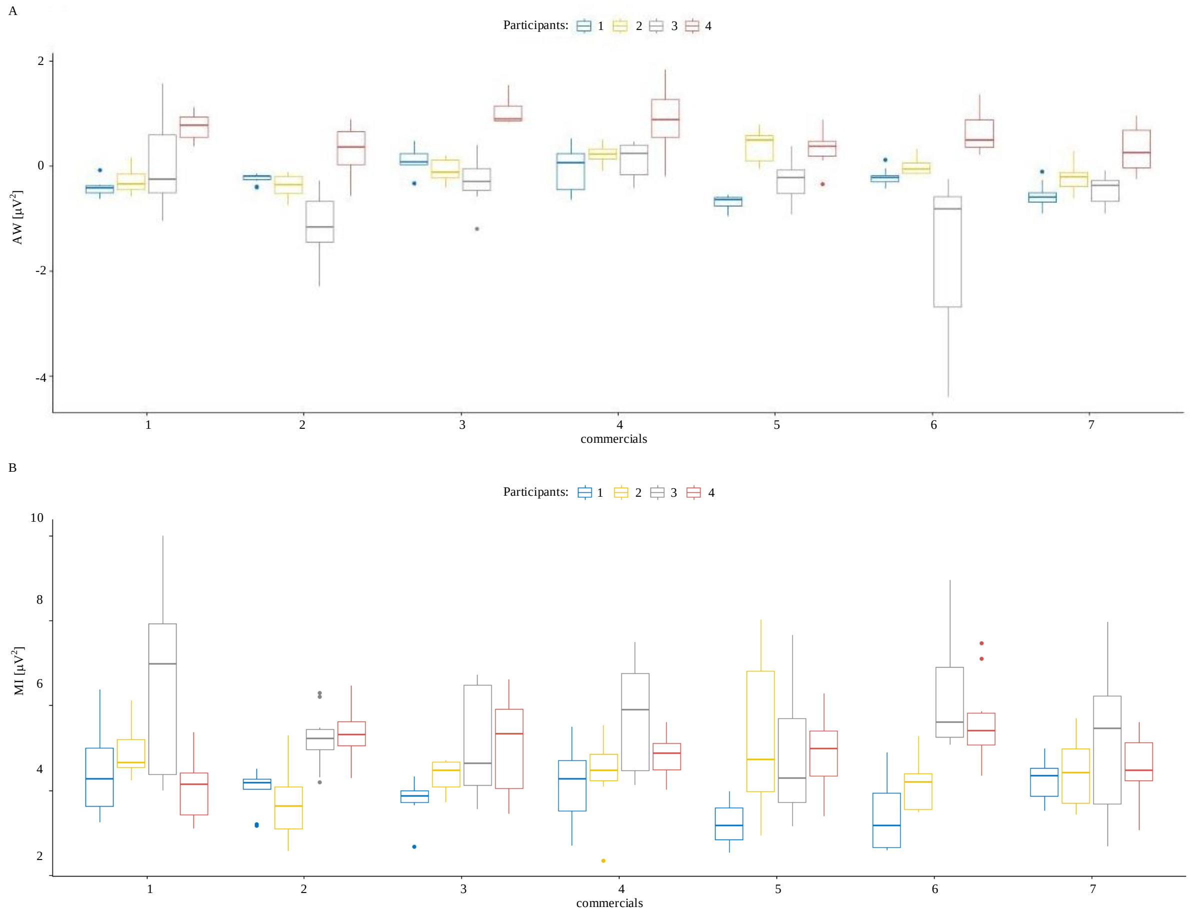


Fig. 3. A) Approach-Withdrawal index (*AW*) and B) Memorization index (*MI*) for each commercial on the individual (participant) level (commercial order 1-7 is shown in Table 1).

From Fig. 3, it can be noticed that there is a high inter-subject variability for each commercial that caused no statistically significant difference for *MI* index on the group level as well as a low statistically significant difference for *AW* index.

Finally, Fig. 4 shows the comparison between normalized total scores of *AW*, *MI* and self-assessment (*SA*) for each commercial. It was found that there is no correlation between *SA* and indices. This means that the objective and subconscious criteria concerning attention and memory is necessary for the detailed observation of the advertisement impact on participants. From Fig. 4, it could be noticed that commercials 1 and 4 have high (attention and memory) impact on participants, commercials 3 and 5 are better suited to attract attention while commercial 6 was memorized better. It is important to mention that commercial 3 lasts shorter than others (20 s vs 28-35 s) which may be the reason for low memory. Commercials 2 and 7 have low induction of attention and low induction of memory as well. Regarding results presented in Fig. 4, normalized total scoring seems to be a promising tool for comparison and visualization of the effects of advertisements on participants. Potential benefit from this type of analysis is obvious: for the investor, it is important to have objective and detailed feedback from customers regarding their attention and memory engagement whilst watching commercials.

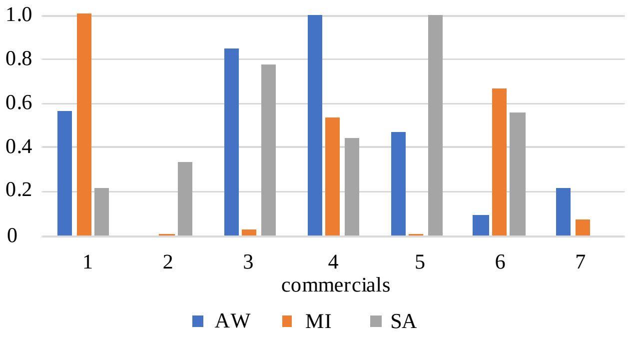


Fig. 4. Comparison of normalized total score for Approach-Withdrawal index (*AW*), Memorization index (*MI*) and Self-assessment (*SA*) for each commercial (commercial order 1-7 is shown in Table 1).

IV. CONCLUSION

In this paper, we have demonstrated a novel approach for the comparison of commercials in the pilot EEG signals based study. Large individual differences in attention and memory involvement were ascertained between participants whilst watching the video content. For that reason, differences in reactions on commercials for each participant were also considered. Individual approach showed the existence of statistically significant differences for attention and memory indices between commercials. All individual scores for *AW*, *MI* and *SA* were summed and normalized giving an effective way for presenting and comparison of different impacts of commercials on participants. However, this type of neuromarketing study can be expanded by simultaneous recordings of heart rate, electrodermal activity and eye tracking activity (along with EEG signals), enabling the assessment of the emotional index and eye tracking features as well. Also, the longer duration of commercials could enable analyzing and comparing individual parts of advertisement, and not just analyzing advertisements as a whole. Future work will be focused on designing a longer experimental paradigm, applying a multimodal approach for data acquisition and advertising differentiation, using multiparametric analysis.

ACKNOWLEDGMENT

This research was supported by the Ministry for Education, Science and Technology Development of Serbia, Serbia.

REFERENCES

[1] A. Bazzani, S. Ravaioli, U. Faraguna & G. Turchetti, "Is EEG Suitable for Marketing Research? A Systematic Review", *Frontiers in Neuroscience*, vol. *14*, pp. *1-16*, 2020.

[2] F. S. Rawnaque, K. M. Rahman, S. F. Anwar, R. Vaidyanathan, T.Chau, F. Sarker, K. A. Al Mamun, "Technological advancements and opportunities in Neuromarketing: a systematic review", *Brain Informatics*, vol. *7*, no. *1*, pp. *1-19*, 2020.

[3] S. G. Meyerding, C. M. Mehlhose, "Can neuromarketing add value to the traditional marketing research? An exemplary experiment with functional near-infrared spectroscopy (fNIRS)", *Journal of Business Research*, vol. *107*, pp. *172-185*, 2020.

[4] M.Y.-T Hsu and J.M.-S. Cheng, "fMRI neuromarketing and consumer learning theory: Word-of-mouth effectiveness after product harm crisis", *European Journal of Marketing*, vol. *52* no. *1/2*, pp. *199-223*, 2018.

[5] A. Hakim, I. Golan, S. Yefet, D. J. Levy, "DeePay: Deep Learning Decodes EEG to Predict Consumer's Willingness to Pay for Neuromarketing", *TechRxiv*. Preprint, 2021.

[6] B. C. Iloka and G. I. Anukwe, "Review of eye-tracking: A neuromarketing technique" *Neuroscience Research Notes*, vol. 3, no. *4*, pp. *29-34*, 2020

[7] M. Zito, A. Fici, M. Bilucaglia, F. S. Ambrogetti and V. Russo, "Assessing the Emotional Response in Social Communication: The Role of Neuromarketing", *Frontiers in Psychology*, vol. *12*, pp. *1784-1797*, 2021.

[8] A. C. Martinez-Levy, D. Rossi, G. Cartocci, M. Mancini, G. D. Flumeri, A. Trettel, F. Babiloni, P. Cherubino, "Message framing, non-conscious perception and effectiveness in non-profit advertising. Contribution by neuromarketing research", *International Review on Public and Nonprofit Marketing*, pp. *1-23*, 2021.

[9] W. Dimpfel, "Neuromarketing: Neurocode-Tracking in Combination with Eye-Tracking for Quantitative Objective Assessment of TV Commercials", *Journal of Behavioral and Brain Science*, vol. *5* no. *4*, Article ID 55367, pp. *137-147*, 2015.

[10] M. Balconi, B. Stumpo, & F. Leanza, "Advertising, brand and neuromarketing or how consumer brain works," *Neuropsychological Trends*, vol. *16* no. *2*, pp. *15-21*, 2014.

[11] G. Vecchiato, W. Kong, A. Giulio Maglione, D. Wei, "Understanding the impact of TV commercials" *IEEE pulse*, vol. *3* no. *3*, pp. *42-47*, 2012.

[12] G. Vecchiato, J. Toppi, L. Astolfi, F. D. V. Fallani, F. Cincotti, D. Mattia, F. Bez, & F. Babiloni, "Spectral EEG frontal asymmetries correlate with the experienced pleasantness of TV commercial advertisements", *Medical & biological engineering & computing*, vol. *49* no. *5*, pp. *579- 583*, 2011.

[13] L. Astolfi, G. Vecchiato, F. De Vico Fallani, S. Salinari, F. Cincotti, F. Aloise, D. Mattia, M. G. Marciani, L. Bianchi, R. Soranzo, F. Babiloni, "The track of brain activity during the observation of tv commercials with the high-resolution EEG technology," *Computational intelligence and neuroscience*, vol. *2009*, Article ID *652078*, pp. *1-8*, 2009.

[14] B. Kostić, "Quantitative analysis of the impact of advertising on attention and emotions," master thesis, University of Belgrade – School of Electrical Engineering, Belgrade, Serbia, 2018.

[15] https://www.lumiere.rs/ostalo/istrazivanja/srbija-top-10-srpske-reklame-u-2012-nase-malo-istrazivanje/ [in Serbian] (last access in June 2021)

[16] R. J. Davidson, "What does the prefrontal cortex "do" in affect: perspectives on frontal EEG asymmetry research," *Biological Psychology*, vol. *67*, no. *1-2*, pp. *219–233*, 2004.

[17] G. Vecchiato, A. G. Maglione, P. Cherubino, B. Wasikowska, A. Wawrzyniak, A. Latuszynska, M. Latuszynska, K. Nermend, I. Graziani, M. R. Leucci, M. Trettel, A. Trettel, "Neurophysiological tools to investigate consumer's gender differences during the observation of TV commercials", *Computational and mathematical methods in medicine*, vol. *2014*, Article ID 912981, pp. *1-12*, 2014.

[18] A. Delorme, S. Makeig, "EEGLAB: an open-source toolbox for analysis of single-trial EEG dynamics", *Journal of Neuroscience Methods*, vol. *134*, pp. *9-21*, 2004.

[19] RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Open-source tool for 3D segmentation and rendering of abdominal CT scans

Katarina Milićević, Otaš Durutović and Milica Janković, *Member, IEEE*

***Abstract*—3D visualization of the size, shape, and location of the kidney stone as well as of the anatomical characteristics of the renal collecting system, surrounding tissue and blood vessels could significantly facilitate the surgeon's treatment planning for urolithiasis. Standard clinical Computed Tomography (CT) software does not offer the flexibility in the 3D display of individual or combined renal phases. In this paper we present a flexible and interactive open-source application for segmentation and 3D visualization of abdominal CT scans for the urolithiasis treatment planning. The usage of the new tool is demonstrated through the clinical examples and its advantages are explained in comparison with the output of the dedicated clinical software.**

***Index Terms*—abdominal CT, minimally invasive surgery, open-source, 3D rendering, 3D segmentation.**

I. Introduction

THE accurate diagnosis is a prerequisite for the effective treatment [1]. In the era of modern medicine, this assumption becomes more important, since the backbone of quality diagnostics consists of performing radiological procedures, primarily computed tomography (CT) imaging, but also radiography, ultrasound, magnetic resonance imaging (MRI), as well as interventional radiological procedures. Radiological procedures are associated with the patient's exposure to ionizing radiation, and therefore it is necessary to use these procedures rationally and optimally.

Low-dose abdominal CT imaging has high sensitivity, specificity, and accuracy for the detection of urolithiasis (kidney stone disease) [2]. Urolithiasis could be treated conservatively (by pain control, hydration, and medical expulsive therapy), surgically (endoscopic methods: ureteroscopy and percutaneous nephrolithotomy) or by extracorporeal shockwave lithotripsy (shockwaves applied outside the body to break the stone) [3,4]. The size, shape, localization, and structure of the kidney stone determine the decision about the treatment method.

The treatment of choice in case of large stones (>2 cm) or complex anatomic factors is percutaneous nephrolithotomy (PCNL) [5-7]. Each percutaneous (through the skin) path to the kidney and the stone implies the passing of instrument through the renal highly vascularized parenchyma, so the optimization of each puncture, even individual, and kidney access is an imperative for the minimization of potential complications. The prerequisite for the minimally invasive surgery is the detailed 3D visualization of the target organ and its anatomic environment [8].

A reliable anatomical 3D visualization is based on the quality segmentation process. One of the directions for abdominal segmentation is by using so called “atlas”, structures that combine the location and shape of anatomical structures and spatial relationships among them [9]. The main obstacle in “atlas” approaches is huge anatomical diversity. Recently, promising results in abdominal segmentation have been achieved using deep learning approaches [10-13], but these methods require large homogeny datasets. However, the greatest accuracy for abdominal organ segmentation has been obtained by multi-level hierarchical strategy combined with neural network approach [14], so in this paper we have used a hierarchical strategy for organ segmentation [15].

Dedicated clinical CT software usually presents independently stone scans, then the structure of the kidney itself, and finally the anatomy of the excretory system. These tools are not interactive, and they do not offer options for controlling the overall process of segmentation and 3D visualization. In this paper, we propose a flexible and interactive open-source tool, so called *3D Gastro CT tool*, for segmentation and 3D rendering of abdominal CT scans that could facilitate the urolithiasis treatment planning. In Section II we have presented the flowchart of the developed application, image preprocessing, segmentation and rendering methodology. Examples of the tool usage in patients with and without kidney stone are illustrated and compared with the output of the dedicated clinical software in Section III. Finally, the conclusion and plans for future work are given in Section IV.

II. The Method

3D Gastro CT tool was developed in Python language using the following libraries and toolkits: matplotlib [16], ndimage [17], SimpleITK (abstraction layer and wrapper around Insight Segmentation and Registration Toolkit, ITK) [18-20] and VTK (Visualization Toolkit) [21]. The graphical user interface (GUI) was developed using PyQt5 bindings for Qt v5 [22]. The code is available at Github link: https://github.com/milicevickatarina/3D-Gastro-CT.

Katarina Milićević is with the University of Belgrade - School of Electrical Engineering, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mail: katarinamilicevic@gmail.com).

Otaš Durutović is with the School of Medicine, University of Belgrade, Department of Surgery and Anaesthesiology, Dr Subotića 8, Belgrade, Serbia and Clinical Centre of Serbia, Urology Clinic, Department of Urology, Resavska 51, Belgrade, Serbia (e-mail: odurutovic@gmail.com).

Milica Janković is with the University of Belgrade - School of Electrical Engineering, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mail: piperski@etf.rs).

The flowchart of the developed application is shown in Fig. 1.

A. *Data management*

3D Gastro CT tool offers the option for reading and viewing a variety of 2D or 3D formats that are supported by SimpleITK reader (*Dicom*, *MetaImage* etc.). This allows the user to select the work directory with images and to explore available image files determining which files are appropriate for further processing.

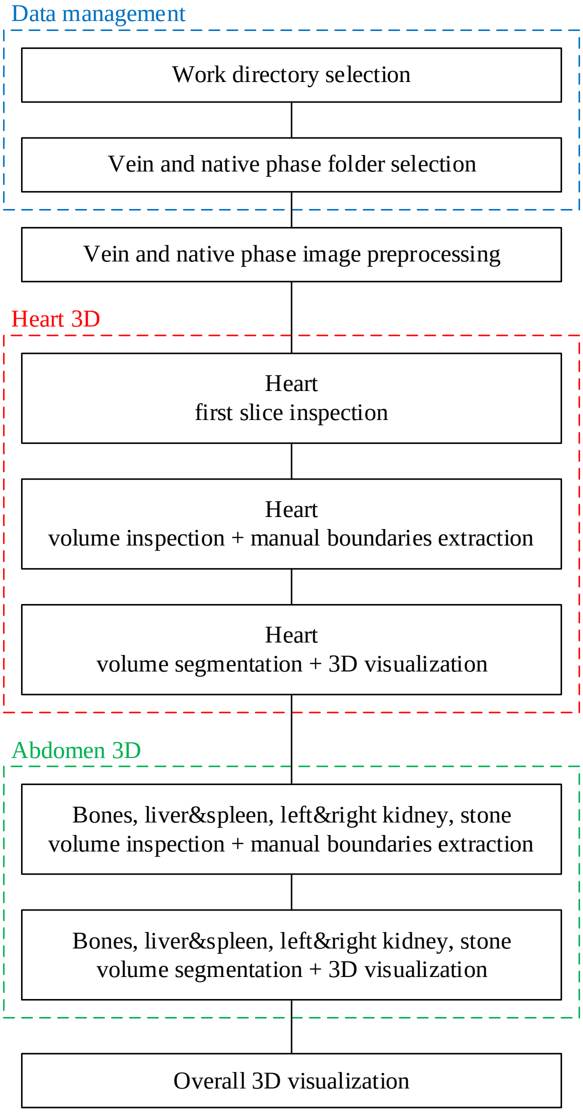


Fig. 1. Flowchart of the 3D Gastro CT tool

For the next step, "Image preprocessing", the selection of folders with native phase images (abdominal images recorded before CT contrast injection) and vein phase images (abdominal images recorded 70-80 s after the CT contrast injection) is necessary. The reason of using these two phases is that grey levels differ well: 1) in vein phase for heart, abdominal organs (liver, spleen, kidney) and stone 2) in native phase for skeleton (unlike vein phase where it overlaps with the intensity of the blood vessels).

Once performed, all preprocessing, segmentation and 3D rendering results would be stored for the future image retrieval. Export option in .stl and .jpg formats of individual and merged rendering results is available.

3D Gastro CT tool was tested using the dataset available from the routine CT imaging (Aquilion Prime CT Scanner, Toshiba Medical Systems, USA) performed for patients with suspected urinary tract stones in the Clinical Center of Serbia. The resolution of CT scans was 512x512 pixels and 588 CT slices in axial projection were available.

B. *Image preprocessing*

The images from the vein and native phases are preprocessed completely automatically using the following steps: 1) linear intensity transfer function for the band [-548, 800] to [0,255], 2) resampling images from 512x512 resolution to 256x256 to reduce time and space resources in further steps, 3) median filtering for noise reduction. Native phase was coregistered with the vein phase using the *SimpleITK* class *ImageRegistrationMethod* (similarity metric: mutual information (Mattes MI), interpolator: sitkLinear, optimizer: gradient descent), [23].

C. *3D segmentation*

The segmentation process is executed hierarchically, organ by organ, on the preprocessed images. The segmentation order is as follows: heart, bones, liver and spleen, kidneys and kidney stone. The segmentation of bones and kidney stone can be performed even if segmentation of previous organs is not finished. The result of individual organ segmentation is displayed rendered in the pop-up interactive window (rotation and zoom options enabled) so the user can decide to save or repeat the segmentation procedure.

C1. *Heart 3D segmentation*

Heart 3D segmentation performs on vein phases images. This procedure includes the following steps, Fig. 2:

- manual selection of heart rectangle region of interest (ROI) boundaries (left, right, upper, lower border) on the first CT slice, Fig. 2A
- fully automatic heart segmentation of the first CT slice using the following steps: 1) gaussian filtering (σ=2), Fig. 2B, 2) image binarization (threshold is set to 10 % of maximal pixel intensity), Fig. 2C, 3) the cross section between the heart rectangle ROI boundaries and the binary image, Fig. 2D, 4) filling holes and removing particles, Fig. 2E.
- fully automatic heart segmentation of the rest of CT slices containing heart using the following steps: 1) gaussian filtering (σ=2), 2) image binarization (threshold is set to 10 % of maximal pixel intensity), 3) the cross section of the binary image, heart rectangle ROI boundaries, and the dilated extracted heart ROI from the previous CT slice (dilation was used to compensate heart movements and different heart surface on successive CT slices)
- opening and closing of the extracted heart volume (cube structural element with the dimension 10).

C2. *Other organs' 3D segmentation*

Bones 3D segmentation performs on native phases images. The procedure is the same as in case of liver&spleen and stone 3D segmentation. Liver&spleen and stone 3D segmentation performs on vein phases images. This procedure includes the following steps:

- histogram presentation for the whole CT volume
- manual selection of upper and lower threshold that correspond to the VOI intensity boundaries on the histogram for image binarization as it is defined in Fig. 3A, B, C (left)
- volume binarization performs on the volume where the heart and bones VOIs are removed for liver&spleen segmentation
- opening and closing of the extracted volume (cube structural element with the dimension 3).

Kidney 3D segmentation performs on vein phases images. This procedure includes the following steps, Fig. 3D:

- manual selection of kidney VOI boundaries (top and bottom slice number, left, right, upper, lower border) that cover the kidney VOI in all CT slices
- histogram presentation for the kidney VOI where bones and liver&spleen volumes are previously removed
- manual selection of upper and lower threshold that correspond to the kidney intensity boundaries on the histogram for volume binarization as it is defined in Fig. 3D (left)
- the cross section between the kidney VOI boundaries and the binary image
- opening and closing of the extracted kidney volume (cube structural element with the dimension 4).

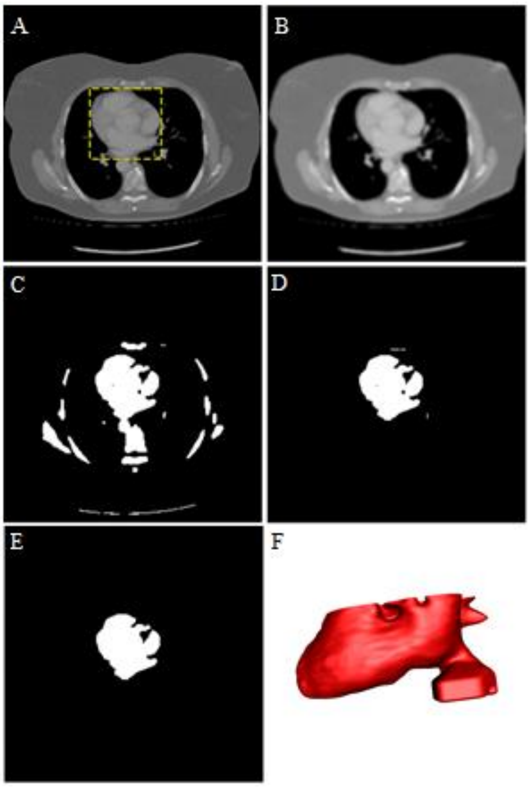


Fig. 2. Heart 3D segmentation process: A) manual selection of heart rectangle ROI boundaries, B) gaussian filtering, C) image binarization, D) cross section with rectangle ROI, E) filling holes and removing particles, F) 3D heart rendering result

D. 3D visualization

3D rendering was performed using the *Marching Cubes* algorithm [24] using the implementation presented in [25]. Segmented images are archived in *.mhd* format compatible with the VTK reader. Individual segmentation results are merged to perform the overall 3D visualization.

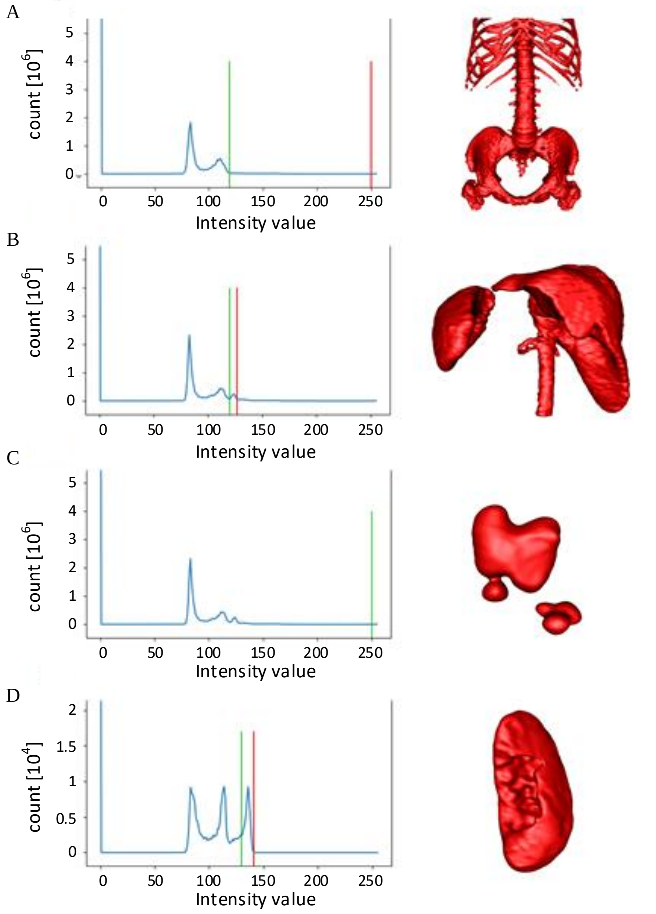


Fig. 3. Bones (A), liver&spleen (B), stone (C), and kidney (D) upper and lower threshold selection on corresponding histograms (left) and individual 3D rendering results (right)

III. Results

An example of the GUI for semi-automated left and right kidney segmentation is shown in Fig. 4. Numerical controls are used for setting values of VOI boundaries. Histogram display is used for determination of lower and upper thresholds. Pop-up windows are used for checking selected kidney boundaries and the display of individual kidney 3D rendering results.

Fig. 5 presents the comparison of the 3D visualization results from *3D Gastro CT tool* and dedicated clinical software for two subjects: the first one is without kidney stone and the second one is with the left kidney stone. For the second subject, the displays are without showing liver and spleen volumes. It could be observed that *3D Gastro CT tool* offers more selectivity in the display where disturbing blood vessels can be completely removed.

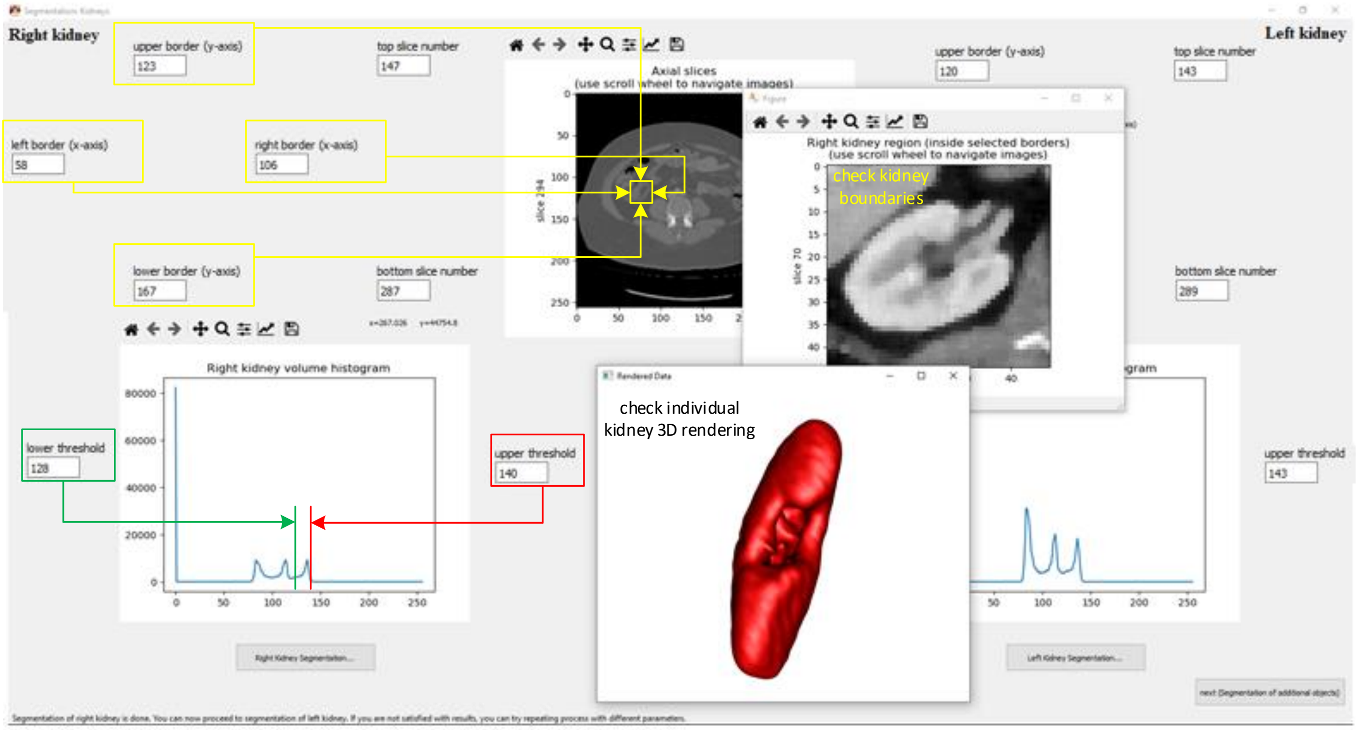


Fig. 4. An example of GUI for setting boundaries and thresholds for the left and right kidney segmentation process, checking boundaries and individual kidney 3D rendering

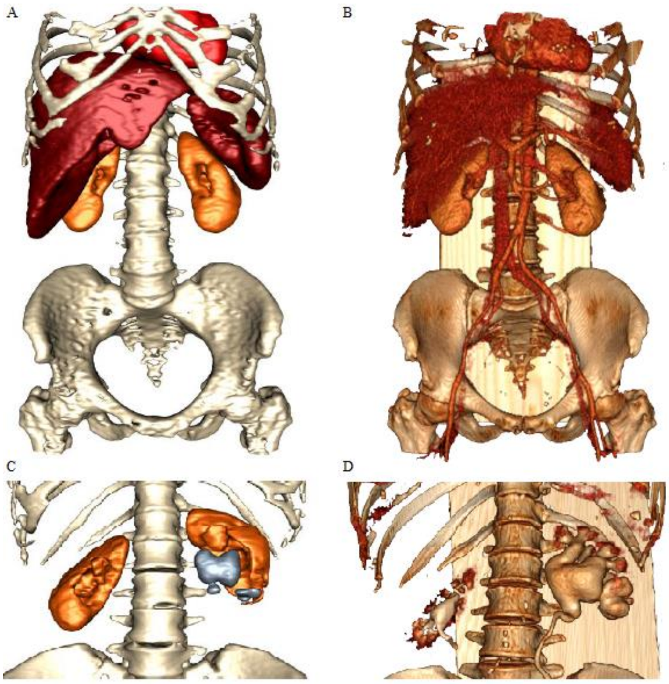


Fig. 5. The comparison of the 3D visualization results from *3D Gastro CT tool* (left) and dedicated clinical software (right) for the patient without (up) and with (down) kidney stone

IV. Conclusion

An open-source semi-automated solution for the visualization of abdomen in patients with urolithiasis is presented in the paper. The segmentation process of individual organs is sequential, and necessary inputs for the realized *3D Gastro CT tool* are native and vein phase CT slices of abdomen. The advantages of the developed tool are demonstrated through examples that are compared with the output from the dedicated clinical software. The usage of the tool is presented on the specific type of urologic CT study, but it is flexible to expand to a wider range of applications. Future work will be focused on testing of the tool on a larger clinical dataset and improving it by the completely automatic algorithm for segmentation.

Acknowledgment

This research was supported by the Ministry for Education, Science and Technology Development of Serbia, Belgrade, Serbia. The authors are especially grateful to dr Milica Stojadinović on the interpretation of CT scans.

References

[1] E. P. Balogh, B. T. Miller, J. R. Ball, *Improving diagnosis in health care*, Washington, USA: National Academies Press, 2015.

[2] J. M. Weinrich, P. Bannas, M. Regier, S. Keller, L. Kluth, G. Adam, F. O. Henes, „Low-dose CT for evaluation of suspected urolithiasis: diagnostic yield for assessment of alternative diagnoses", *American Journal of Roentgenology*, vol. *210*, no. *3*, pp. *557-563*, 2018.

[3] L. Tzelves, C. Türk, A. Skolarikos, „European Association of Urology Urolithiasis Guidelines: Where Are We Going?", *European Urology Focus*, vol. *7*, no. *1*, pp. *34–38*, 2021.

[4] C. Türk, A. Neisius, A. Petrik, A. Seitz, A. Skolarikos, K. Thomas K, G. Gambaro, „Urolithiasis", European Association of Urology Guidelines, Available from: https://uroweb.org/guideline/urolithiasis/ (last access June 2021)

[5] D. Assimos, A. Krambeck, N. L. Miller, M. Monga, M. H. Murad, C. P. Nelson, K.T. Pace, V. M. Pais, M. S. Pearle, G. M. Preminger, H. Razvi, O. Shah, B. R. Matlaga, „Surgical Management of Stones: AUA/Endourology Society Guideline", Available from: https://www.auanet.org/guidelines/kidney-stones-surgical-management-guideline#x3158 (last access June 2021)

[6] B. Baralo, P. Samson, D. Hoenig, A. Smith, „Percutaneous kidney stone surgery and radiation exposure: A review", *Asian journal of urology*, vol. *7*, no. *1*, pp. *10–17*, 2020.

[7] O. Durutović, „Evaluation of percutaneous nephrolithotomy efficacy in the treatment of patients with kidney stone", PhD thesis, University of Belgrade, School of Medicine, Belgrade, Serbia, 2016.

[8] A. C. Westphalen, R. Y. Hsia, J. H. Maselli, R. Wang, R. Gonzales, „Radiological imaging of patients with suspected urinary tract stones: national trends, diagnoses, and predictors", *Academic emergency medicine : official journal of the Society for Academic Emergency Medicine*, vol. *18*, no. *7*, pp. *699–707*, 2011.

[9] B. Oliveira, S. Queirós, P. Morais, H. R. Torres, J. Gomes-Fonseca, J. C. Fonseca, J. L. Vilaça. „A novel multi-atlas strategy with dense deformation field reconstruction for abdominal and thoracic multi-organ segmentation from computed tomography", *Medical Image Analysis*, vol. *45*, pp. *108-120*, 2018.

[10] O. Ronneberger, P. Fischer, T. Brox T, „U-Net: Convolutional Networks for Biomedical Image Segmentation", In: Navab N., Hornegger J., Wells W., Frangi A. (eds) Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol. *9351*, pp. *234-241*, Springer, Cham, 2015.

[11] F. Milletari, N. Navab, S. A. Ahmadi, „V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation", Fourth International Conference on 3D Vision (3DV), pp. *565-571*, Stanford, CA, 2016.

[12] P. Hu, F. Wu, J. Peng, Y. Bao, F. Chen, D. Kong, „Automatic abdominal multi-organ segmentation using deep convolutional neural network and time-implicit level sets", *International Journal of Computer Assisted Radiology and Surgery*, vol. *12*, no. *3*, pp. *399–411*, 2017.

[13] E. Gibson, F. Giganti, Y. Hu, E. Bonmati, S. Bandula, K. Gurusamy, B. Davidson, S. P. Pereira, M. J. Klarkson, D. C. Barratt, „Automatic Multi-Organ Segmentation on Abdominal CT With Dense V-Networks", *IEEE Transactions on Medical Imaging*, vol. *37*, no. *8*, pp. *1822-1834*, 2018.

[14] M. A. Selver, „Segmentation of abdominal organs from CT using a multi-level, hierarchical neural network strategy", *Computer Methods and Programs in Biomedicine*, vol. *113*, no. *3*, pp. *830-852*, 2014.

[15] K. Milićević, „3D segmentation and visualization of abdominal computed tomography scans", BSc. Thesis, University of Belgrade – School of Electrical Engineering, Belgrade, Serbia, 2020.

[16] J. D. Hunter, „Matplotlib: A 2D Graphics Environment", *Computing in Science & Engineering*, vol. *9*, no. *3*, pp. *90-95*, 2007.

[17] https://docs.scipy.org/doc/scipy/reference/ndimage.html (last access June 2021)

[18] W. Schroeder, K. Martin, B. Lorensen, *The Visualization Toolkit*, 4th ed, New York, USA: Kitware, 2006.

[19] R. Beare, B. C. Lowekamp, Z. Yaniv, „Image Segmentation, Registration and Characterization in R with SimpleITK", *J Stat Softw*, vol. *86*, no. *8*, pp. *6595*, 2018.

[20] Z. Yaniv, B. C. Lowekamp, H. J. Johnson, R. Beare, „SimpleITK Image-Analysis Notebooks: a Collaborative Environment for Education and Reproducible Research", *J Digit Imaging*, vol. *31*, no. *3*, pp. *290-303*, 2018.

[21] B. C. Lowekamp, D. T. Chen, L. Ibáñez, D. Blezek, „The Design of SimpleITK", *Front. Neuroinform*, vol. *7*, no. *45*, pp. 1-14, 2013.

[22] https://pypi.org/project/PyQt5/ (last access June 2021)

[23] https://github.com/InsightSoftwareConsortium/SimpleITK-Notebooks/blob/master/Python/60_Registration_Introduction.ipynb (last access June 2021)

[24] W. E. Lorensen, H. E. Cline, „Marching cubes: A high resolution 3D surface construction algorithm", *ACM siggraph computer graphics*, vol. *21*, no. *4*, pp. *163-169*, 1987.

[25] https://github.com/lorensen/VTKExamples (last access June 2021)

ЕЛЕКТРОЕНЕРГЕТИКА
/
POWER ENGINEERING

(EE/EEI)

Metod za inženjersku procenu proizvodnje vetroelektrane

Vladimir A. Katić, *Senior Member, IEEE*

***Abstrakt* — U ovom radu je razmatrana mogućnost brzog i jednostavnog procenjivanja proizvodnje električne energije planirane vetroelektrane na osnovu korišćenja raspoloživih podataka. Pošlo se od pretpostavke da su raspoložive usrednjene godišnje vrednosti brzine i gustine snage energije vetra. Razmatranjem različitih opcija za procenjivanje proizvodnje, predložena je jedna metoda. Radi verifikacije ona je primenjena za jednu lokaciju i četiri vrste vetrogeneratora. Metoda je verifikovana poređenjem sa izračunatom energijom na bazi stvarnih rezultata merenja brzine vetra. Postignuta je zadovoljavajuća tačnost pogodna za inženjerske procene, čime je metoda opravdala mogućnost primene.**

***Ključne reči* — Vetroenergija, Vetroelektrane, Metod procene proizvodnje.**

I. Uvod

Primena obnovljivih izvora energije danas je prihvaćena kao jedno od rešenja za zamenu fosilnih goriva, čije rezerve se ubrzano iscrpljuju. Istovremeno, time se omogućuje smanjenje emisije CO_2 i drugih gasova staklene bašte, za koje je pokazano da doprinose povećanju prosečne temperature na Zemlji i time dovode do klimatskih promena [1].

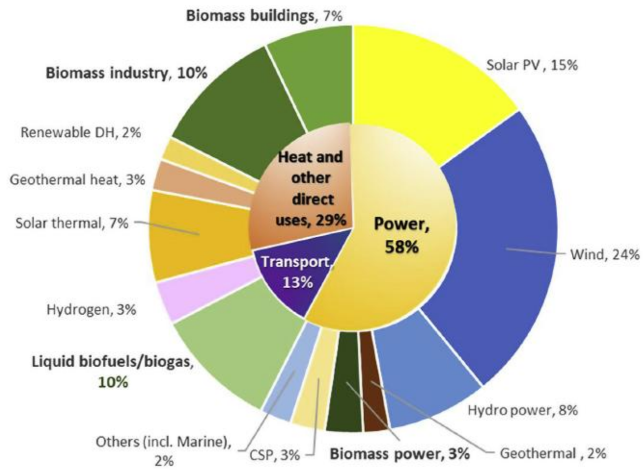
Za generisanje električne energije iz obnovljivih izvora najčešće se koriste energije vetra, vode, sunca, biomase i geotermalna. Trenutno su u svetu najveći instalisani kapaciteti u vetroelektranama sa dobrim perspektivama rasta. Na sl. 1 prikazana je struktura primene obnovljivih izvora prema planu za 2050. god., gde se vidi da će najveći deo obnovljivih izvora biti angažovan za dobijanje električne energije (58%), od čega 41% (ili 24% od ukupnog) će činiti vetroelektrane [2].

U Srbiji je utvrđen značajan potencijal za korišćenje energije vetra i već su izgrađene ili su u izgradnji vetroelektrane ukupne snage od 566 MW [3,4]. Strategijom razvoja energetike Srbije predviđeno je da se u narednom periodu (do 2030. god.) izgradi ukupno 600 MW, ali su procenjeni potencijali znatno veći (1300 MW, pa i više) [4,5]. To ukazuje na potrebu da se kontinuirano istražuju adekvatne lokacije i mogućnosti izgradnje novih vetroelektrana.

Jedan od prvih preduslova za određivanje pogodne lokacije je poznavanje energetskog potencijala. Ovaj potencijal se određuje preko brzine vetra i gustine snage vetra u određenoj oblasti ili na mikrolokaciji. U svetu je poznat veliki broj studija, naučnih radova i drugih publikacija na ovu temu. Pored toga postoji i veći broj web-sajtova, preko kojih se može dobiti procena vetro-energetskog potencijala, a jedan od takvih je i „Global Wind Atlas“ (https://globalwindatlas.info/) [6]. U Srbiji je rađeno nekoliko studija, a jedan od primera je i studija „Atlas vetrova AP Vojvodine“ [3].

Međutim, problem je što su u tim publikacijama ili na sajtovima date uprosečene godišnje vrednosti brzine i gustine energije vetra. Razlog za to je velika promenljivost energije vetra, kao i zavisnost od mnogih lokalnih faktora (mikroklime, reljefa, orografije, i dr.), što sve umanjuje praktičnu primenljivost ovih podataka. Da bi se dobile tačnije procene za odabranu lokaciju, preporučuju se dugotrajna, specijalizovana merenja parametara energije vetra u trajanju od bar godinu [7]. Na taj način postavljaju se pouzdane osnove na bazi kojih se može dovoljno tačno odrediti količina električne energije, koju bi planirana vetroelektrana proizvela na datoj lokaciji. Ipak, ovakva merenja su skupa, dugo traju i zahtevaju angažovanje specijalizovanih stručnjaka, eksperta u ovoj oblasti.

U ovom radu se razmatraja mogućnost brzog i jednostavnog procenjivanja proizvodnje električne energije planirane vetroelektrane na osnovu korišćenja raspoloživih podataka. Cilj je da se korišćenjem raspoloživih podataka srednje godišnje vrednosti brzine i gustine snage energije vetra, kao i karakteristika vetrogeneratora razvije adekvatna metoda za inženjersku procenu proizvodnje jedne vetroelektrane. Razmatranjem različitih opcija za procenjivanje proizvodnje, predložena je metoda. Radi verifikacije ona je primenjena za jednu lokaciju i četiri vrste vetrogeneratora. Metoda je verifikovana poređenjem sa izračunatom energijom na bazi stvarnih rezultata merenja brzine vetra.



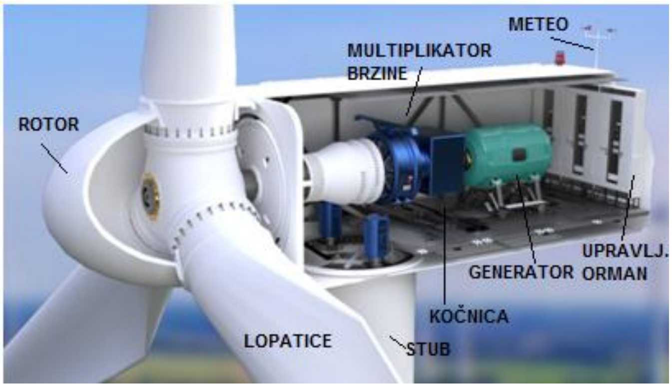
Sl. 1. Plan korišćenja obnovljivih izvora za 2050. god. [2]

II. Način pretvaranja energije vetra

Energija vetra, kao linearno vazdušno kretanje pretvara se u korisnu rotacionu mehaničku energiju posredstvom trokrake

Vladimir A. Katić je sa Fakulteta tehničkih nauka, Univerziteta u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: katav@uns.ac.rs).

vetroturbine. Vetroturbina je postavljena na horizontalnu osovinu na vrhu nosećeg stuba. Zbog male brzine obrtanja (20-30 min^{-1}) ova osovina se naziva „spora“. Da bi se postigna potrebna ugaona brzina za pogon električnog generatora koristi se mehanički sklop kojim se povećava brzina - multiplikator brzine. Njime se pogon prenosi na „brzu“ osovinu, koja je direktno spregnuta na električni generator. Na njoj se nalazi i mehanička kočnica, za zaustavljanje pogona u slučajevima velikih brzina vetra ili havarije. Generator pretvara mehaničku energiju prenetu sa vetroturbine u električnu energiju, koja predstavlja finalni proizvod rada vetrogeneratora. Ta električna energija se zatim prilagođava parametrima mreže putem energetskih elektronskih pretvarača (obično dvostrukih pretvarača tipa AC/DC–DC/AC), a naponski pomoću energetskih transformatora. Na sl. 2 prikazan je presek vetrogeneratora, odnosno princip vetro-električnog agregata sa nekim od pomenutih sklopova.



Sl. 2. *Princip konverzije energije vetra u električnu energiju [8]*

III. PRORAČUN ENERGETSKIH MOGUĆNOSTI

Generalno energija vetra se izražava preko gustine snage, kao:

$$P_v/A = \tfrac{1}{2}\,\rho\cdot v^3, \qquad (1)$$

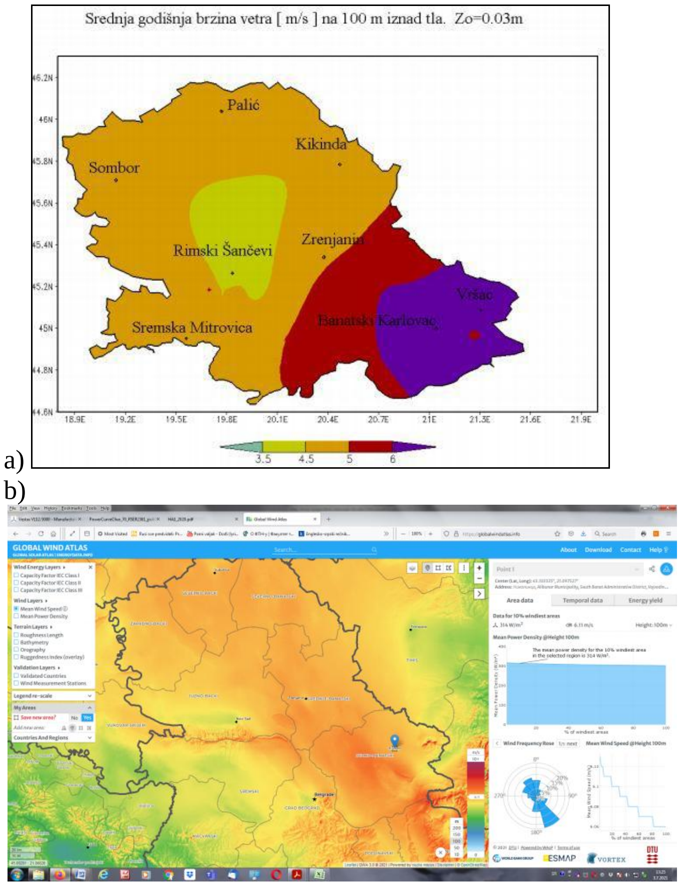
gde je P_v snaga vetra, A površina osnove cilindra kroz koji struji vetar, ρ specifična gustina vazduha na određenoj nadmorskoj visini i v brzina vetra [1,3,7]. Vidi se da je najuticajniji parametar brzina vetra, pa je njeno precizno određivanje od ključnog značaja.

A. Brzina vetra

Brzina vetra dnevno se meri u široko postavljenoj mreži hidro-metoeroloških stanica, kako u Srbiji tako i u svetu. Obično se merenja vrše na 10 m visine na svakih 10 minuta, tj. sa 10-min. usrednjavanjem. Za potrebe vetroenergetike, to nije dovoljno, jer se moderne vetroturbine nalaze na visinama od 100 m ili više. Iz tog razloga, potrebno je ili vršiti merenja na visini vetroturbine ili izmerene vrednosti brzine preračunati na odgovarajuću visinu, uvažavajući i orografiju terena izraženu koeficijentom z_0.

Međutim, iz naučne literature i raspoloživih web sajtova može se dobiti vrednost prosečne godišnje brzine vetra na određenoj lokaciji i visini. Na primer, iz studije “Atlas vetrova AP Vojvodine” može se dobiti prikaz prosečne godišnje brzine vetra na visini od 100 m (sl. 3.a) sa koje se vidi da ona u jugo-istočnoj Vojvodini iznosi preko 5 m/s, odnosno preko 6 m/s [3]. Za isti deo Vojvodine, sa sajta „Global Wind Atlas”, dobija se da je prosečna godišnja brzina vetra na 100 m iznosi od 5,1 m/s do 9,2 m/s [6].

Za projektovanje novih vetroelektrana ključni podatak je mogući iznos proizvedene električne energije, pa je pitanje da li se i na osnovu ovakvih podataka može dobiti prihvatljiva procena o generisanoj energiji.



Sl. 3. P*rosečne godišnje brzine vetra u Vojvodini na 100 m iznad tla prema: a) “Atlasu vetrova APVojvodine” [3] i b) “Global Wind Atlas-u”[6]*

B. Snaga vetra

Za odabir snage vetroturbine u sklopu vetroelektrane potrebno je proračunati snagu vetra na zahvatu vetrogeneratora. Merenja pokazuju da raspodelu brzina najbolje prati *Weibull*-ova funkcija raspodele, koju karakterišu faktor oblika (k) i kaktor skaliranja (c) [1,3,7]. Međutim, za većinu lokacija prihvatljiva je vrednost faktora oblika k=2, kad se ova funkcija naziva *Rayleigh*–ova. Sada se srednja godišnja gustina snage može izračunati kao [7]:

$$P_{vSR}/A = 6/\pi\ \tfrac{1}{2}\ \rho\cdot(v_{SR})^3 \qquad (2)$$

gde je v_{SR} srednja vrednost brzina za određenu lokaciju, koja se dobija iz raspoloživih podataka (kao na sl. 3), a gustina vazduha obično ρ =1,225 kg/m^3 (mada pri većim nadmorskim visinama potrebno ju je posebno proračunati).

Za određen tip vetrogeneratora, može se iz podataka o vetroturbini dobiti površina zahvata, kao $A=R^2\pi$, gde je R dužina lopatica. Na primer, vetrogenerator Vestas V112-3.08 snage 3.08 MW, čije su karakteristike date u tabeli 1 [9], svojim lopaticama od 56 m prebriše površinu od 9.852 m^2, te uz prosečnu godišnju brzinu vetra od v_{SR} =6,12 m/s, prihvata srednju godišnju snagu vetra od 2,8 MW.

C. Snaga vetrogeneratora

Snaga vetra se koristi za pokretanje vetroturbine, čiji rad se može podeliti u četiri zone: 1 – polazna $P_{t1}=0$ ($0<v<v_{in}$), 2 – radna $P_{t2}=f(v)$ ($v_{in}<v<v_n$), 3 - nominalna $P_{t3}=Pn$ ($v_n<v<v_{out}$) i 4 – zaustavna $P_{t4}=0$ ($v>v_{out}$). Na slici 4 prikazane su krive snage nekoliko komercijalnih vetrogeneratora snaga od 2 MW do 3 MW i to REpower MM82 (2 MW), Enercon E 82/2300 (2,3 MW), Nordex N90/2500 (2,5 MW) i Vestas V112/3.08 (3.08 MW). Mogu se uočiti četiri pomenute radne zone.

Za određivanje proizvodnje vetrogeneratora interesantne su radna i nominalna zona, dok je u preostale dve zone snaga turbine jednaka nuli (u polaznoj zoni, jer nije ostvaren dovoljan momenat za pokretanje, a u zaustavnoj zoni da bi se izbegla prevelika mehanička naprezanja). U radnom delu snaga turbine definisana je koeficijentom snage C_p, pa je sada snaga zavisna od brzine vetra na kub:

$$P_{t2} = C_p \cdot P_v = ½ \cdot C_p \cdot \rho \cdot A \cdot v^3 = 3/\pi \cdot C_p \cdot \rho \cdot A \cdot (v_{SR})^3 \tag{3}$$

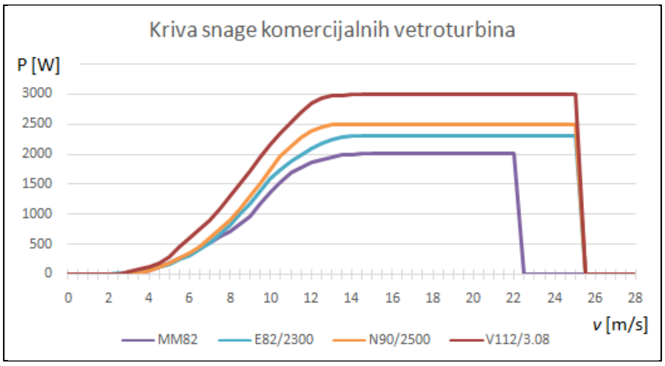
Za proizvodnju najbolje je da turbina radi sa maksimalnom snagom (C_{pmax}), odnosno da je $P_t = C_{pmax} \cdot P_v$. Na primer, za pomenuti model Vestas V112/3080 podaci iz literature definišu C_{pmax}=0,43, pa bi prosečna snaga turbine na lokaciji sa sl. 3.b) bila 43% snage vetra, odnosno 1,2 MW [10].

Međutim, u realnim uslovima ovo nije ispunjeno, jer C_p nije konstantno, već funkcija koeficijenta brzohodnosti λ (*tip-speed ratio*) i ugla zakrenutosti lopatica β [7,10]. Koeficijent brzohodnosti zavisi od brzine, pa kriva snage odstupa od kubnog zakona i ima oblik istegnute „S“ krive (sl. 4). Sa sl. 4 vidi se da je moguće izvršiti linearnu aproksimaciju ovh krivih u radnoj zoni. Primer za vetroturbinu V112/3.08 predstavljen je na sl. 5. Sada se kriva snage vetroturbine može linearizovati, tj. napisati kao:

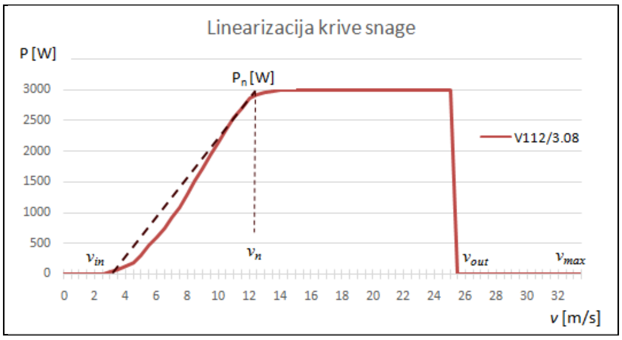
$$P_t = \begin{cases} P_{t1} = P_{t4} = 0 \quad , 0 < v < v_{in} \ \wedge \ v > v_{out} \\ P_{t2} = P_n \cdot \dfrac{v - v_{in}}{v_n - v_{in}} \quad , \quad v_{in} < v < v_n \\ P_{t3} = P_n \quad , \quad v_n < v < v_{out} \end{cases} \tag{4}$$

gde je P_n nominalna snaga u [W], v_{in} upadna brina vetra (*cut-in*), v_n nominalna brzina, v_{out} zaustavna brzina (*cut-out*), a v_{max} maksimalna brzina vetra, sve u [m/s].

Brzina vetra je promenljiva i obično varira u širokim opsezima, Na primer, u jugo-istočnoj Vojvodini (sl. 3) varira u širokom opsegu sa udarima i do 38 m/s [11]. Međutim, meri se sa 10-minutnim usrednjavanjem, čime se nivelišu kratkotrajni udari, pa je mala verovatnoća brzina iznad v_{out}.



Sl. 4. Kriva snage komrcijalnih vetroturbina [9]



Sl. 5. Linearizacija krive snage (primer za vetroturbinu V112/3.08)

Za brze proračune ovakva linearizacija i aproksimacija maksimalne brzine je zgodna, jer se sada srednja snaga turbine u celom rasponu brzina vetra može izraziti kao:

$$P_{tSR} = P_n \cdot \left(1 - \frac{v_{in} + v_n}{2 \cdot v_{out}} \right) \tag{5}$$

Izlazna, električna snaga vetrogeneratora P_g zavisi od koeficijenta efikasnosti mehaničkog prenosa (reduktora) η_{meh} i elektro-mehaničkog pretvaranja (generatora) η_{gen}. Sada se za srednju snagu vetrogeneratora može napisati:

$$P_{gSR} = \eta_{gen} \cdot \eta_{meh} \cdot P_{tSR} \tag{6}$$

Na primer, za vetroturbinu V112-3.08, a uz pretpostavku koeficijenata efikasnosti η_{meh}=0,9 i η_{gen}=0,95, dobija se za srednju snagu vetrogeneratora 1,03 MW.

D. Procena proizvedene električne energije

Najjednostavnija procena dobijene električne energije iz jednog vetrogeneratora može se dobiti, ako se iskoristi izraz (3) i pretpostavi rad sa C_{pmax}. Ona se dobija poznavanjem samo srednje brzine vetra na nekoj lokaciji (v_{SR}) i karakteristika odabranog vetrogeneratora. Sad je, na bazi (6), generisana godišnja električna energija (E_g) data sa:

$$E_g = P_g \cdot 8760 = \eta_{gen} \cdot \eta_{meh} \cdot 3/\pi \cdot C_{pmax} \cdot \rho \cdot A \cdot (v_{SR})^3 \cdot 8760 \text{ [MWh]} \tag{7}$$

Na primer, iz vetroturbine V112-3.08 i na lokaciji sa sl. 3.b), može se godišnje očekivati 8.649,23 MWh. Međutim, tu nije uvažena nelinearna priroda koeficijenta snage C_p, pa ni kriva snage vetroturbine (sl. 4), pa se ovakav metod ne može prihvatiti.

Uvažavanjem ovih faktora i uz predloženu linearizaciju krive snage, te usrednjavanje snage turbine prema (5) i uvrštavanjem u (6) za godišnju proizvodnju jednog vetrogenertora dobija se:

$$E_g = \eta_{gen} \cdot \eta_{meh} P_n \cdot [1-(v_{in}+v_n)/2v_{out}] \cdot 8760 \text{ [MWh]} \tag{8}$$

Na primer, za pomenuti vetrogenerator Vestas V112-3.08 ukupna godišnja proizvodnja bila bi 16.148,01 MWh. Međutim, korišćenje izraza (5) pretpostavlja se ravnomerna raspodela brzina vetra unutar opsega usrednjavanja, što u praksi nije slučaj, pa ni ovaj metod nije prihvatljiv.

Stvarna raspodela brzina vetra prati pomenutu *Rayleigh*–ovu funkciju, pa se pomoću nje može izračunati verovatnoća pojavljivanja brzina većih od neke zadate (v_x). Kako su poznati faktori oblika i skaliranja, dobija se jednostavan izraz za ovu verovatnoću [7,12]:

$$\Psi(v \geq v_x) = e^{-\frac{\pi}{4}\left(\frac{v_x}{v_{SR}}\right)^2} \tag{9}$$

Ovakva forma je pogodana, jer zahteva poznavanje samo srednje vrednosti brzine vetra. Uz linearizaciju krive snage, može se uzeti da je srednja snaga u radnoj zoni jednaka $P_n/2$ (srednja linija trougla), a u nominalnoj zoni P_n. Sada se za srednju godišnju snagu vetroturbine može očekivati da je:

$$P_{tSR} = [1/2 \cdot \Psi(v_{in}<v<v_n)+\Psi(v_n<v<v_{out})] \cdot P_n \qquad (10)$$

Kako je:

$$\Psi(v_{in}<v<v_n)=\Psi(v>v_{in})-\Psi(v>v_n), \qquad (11)$$

odnosno

$$\Psi(v_n<v<v_{out})=\Psi(v>v_n)-\Psi(v>v_{out}) \qquad (12)$$

izraz (10) se može napisati kao:

$$P_{tSR} = [1/2 \cdot \Psi(v>v_{in})+1/2 \cdot \Psi(v>v_n)-\Psi(v>v_{out})] \cdot P_n \qquad (13)$$

S obzirom da se verovatnoća $\Psi(v>v_{out})$ može zanemariti u odnosu na ostale, dobija se:

$$P_{tSR} = [\Psi(v>v_{in})+\Psi(v>v_n)] \cdot P_n/2 \qquad (14)$$

Sad se ukupna godišnja proizvodnja električne energije nekog vetrogeneratora može proceniti na:

$$E_g = \eta_{gen} \cdot \eta_{meh} \cdot [\Psi(v>v_{in})+\Psi(v>v_n)] \cdot P_n/2 \cdot 8760 \text{ [MWh]} \qquad (15)$$

Na primer, za srednju godišnju brzinu v_{SR}=6,12 [m/s] i vetrogenerator V112-3.08, godišnji prinos električne energije može se proceniti na 9.951,8 MWh.

IV. Provera tačnosti metode

Radi provere tačnosti predložene metode procene godišnje proizvodnje električne energije jednog vetrogeneratora biće iskorišćeni stvarni rezultati merenja brzine vetra na visini 60 m (sa usrednjavanjem na 10 minuta) na lokaciji u Banatu [13]. Za poređenje su odabrana četiri različita tipa vetrogeneratora snaga oko 3 MW, čije osnovne karakteristike su date u tabeli 1.

Rezultati su predstavljeni u tabeli 2. Za sva četiri odabrana vetrogeneratora izračunate su vrednosti proizvedene energije u jednogodišnjem period (E_g) prema jednačinama (7), (8) i jednačini (15) - predloženom metodu. Radi tačnog poređenja, rezultati merenja brzine vetra preračunati su na visinu stuba odabranog vetrogeneratora. Na kraju je izračunata greška predložene metode u odnosu na izmereni rezultat.

Može se uočiti da se metoda može primenjivati za brze procene generisane električne energije nekog vetrogeneratora, jer je greška procene u svim slučajevima ispod 5%. Takođe, vidi se da korišćenje jednačina (7) i (8) nije pogodno, jer je greška značajno veća, tj. neprihvatljiva.

V. Zaključak

Korišćenje energije vetra za proizvodnju električne energije beleži sve veći porast i otkriva u sve većoj meri svoj pravi potencijal i značaj. U radu je predložena metoda, kojom se korišćenjem raspoloživih podataka o srednjoj godišnjoj vrednosti brzine i gustine snage energije vetra, kao i karakteristika vetrogeneratora može proceniti proizvodnja električne energije jedne vetroelektrane. Metoda je testirana i prikazala je zadovoljavajuću tačnost, te se može pouzdano primeniti za brze inženjerske procene rada vetroelektrane.

Tabela 1. Osnovne karakteristike odabranih vetrogeneratora [9]

Vetrogenerator	**V112-3.08**	**SWT-3.0-101**	**N131/3000**	**GE 3.2-130**
Proizvođač	Vestas	Siemens	Nordex	Gen.Electric
Snaga (MW)	3.08	3.0	3.0	3.2
v_{in} (m/s)	3,0	3,0	3,0	2,0
v_n (m/s)	12,0	12,5	11,5	12,0
v_{out} (m/s)	25,0	25,0	20,0	25,0
Lopatice, R(m)	56	50,5	65,5	65
Visina (m)	94,0	115	114	110

Tabela 2. Poređenje rezultata primenjene metode i merenja na različitim vrstama vetrogeneratorima

Vetrogenerator /Metod (v_{SR}=6,12 m/s)	**V112-3.08** E_g [MWh]	**SWT-3.0-101** E_g [MWh]	**N131/3000** E_g [MWh]	**GE 3.2-130** E_g [MWh]
Jedn. (7)	8.649,2	7.573,0	12.699,5	12.344.5
Jedn. (8)	16.148,0	15.503,9	14.324,2	17.256,5
Jedn. (15) (Predl.metod)	10.048,1	9.823,3	10.109,8	11.657,9
Merenje	9.951,8	9.619,9	10.528,4	12.085,8
Greška metode	0,97%	2,11%	3.98%	3,54%

Literatura

[1] L. Freris, D. Infield, *„Renewable Energy in Power Systems“*, J. Wiley & Sons, Chichester, 2008.

[2] D. Gielen et al., „The role of renewable energy in the global energy transformation“, *Energy Strategy Reviews*, Vol. 24, 2019, pp.38-50.

[3] V.A. Katić i sarad., *“Atlas vetrova AP Vojvodine“*, Studija, Pokrajinski sekretarijat za energetiku i mineralne sirovine, Fakultet tehničkih nauka, Novi Sad, 2008

[4] https://mre.gov.rs

[5] R. Putnik i sarad., *„Mogućnost korišćenja energije vetra za proizvodnju električne energije“*, Studija, Elektroprivreda Srbije, Beograd, 2002.

[6] https://globalwindatlas.info/

[7] Ž. Đurišić, *„Vjetroelektrane“*, Akademska misao, Beograd, 2019.

[8] https://www.infineon.com/cms/en/applications/industrial/wind-energy-systems/

[9] https://www.thewindpower.net

[10] C. Carrillo et al., “Review of power curve modelling for wind turbines”, *Renewable and Sustainable Energy Reviews*, Vol.21, 2013, pp.572-581.

[11] P. Katić i dr. *„Klima SAP Vojvodine“*, Poljopr. fakultet, Novi Sad, 1979.

[12] Z. Ivanović, V.A. Katić, *„Obnovljivi izvori električne energije – vežbe“*, Fakultet tehničkih nauka, Edicija Tehničke nauke - udžbenici, br.707, Novi Sad, 2018.

[13] M. Hrnjak, *“Glavne karakteristike i komponente vetroelektrane u jugoistočnom Banatu”*, Master rad, Fakultet tehničkih nauka, Novi Sad, 2021.

Abstract

In this paper, the possibility of quick and simple estimation of electricity production of the planned wind power plant based on the use of available data is considered. We started from the assumption that the average annual values of wind speed and power density are available. By considering different options for estimating production, one method has been proposed. For verification, it was applied to one location and four types of wind turbines. The method was verified by comparison with the calculated energy based on the actual results of wind speed measurements. Satisfactory accuracy suitable for engineering assessments was achieved, which justified the method of application.

A Method for Engineering Assessment of Wind Farm Production

Vladimir A. Katić

Modelovanje sistema za regulaciju pobude sinhronog generatora primjenom nelinearnog ARX modela

Mihailo Micev, *Student Member, IEEE*, Martin Ćalasan i Milovan Radulović, *Member, IEEE*

***Apstrakt*—U ovom radu prikazana je primjena nelinearnog auto – regresionog modela sa spoljašnjim ulazom (ARX) u cilju modelovanja sinhronog generatora tokom rada sa opterećenjem. Ulazni signal u sinhroni generator predstavljen je naponom pobude, dok je izlazni signal ustvari napon na krajevima generatora. Stoga je za estimaciju parametara usvojenog modela neophodno koristiti samo napon pobude i izlazni napon generatora, koji predstavljaju ulazno – izlazni set podataka. Dodatno, predložena identifikaciona procedura se bazira na promjeni referentnog napona sistema za regulaciju pobude sinhronog generatora. Opisana procedura je implementirana u programskom paketu MATLAB Simulink. Rezultati dobijeni korišćenjem nelinearnog ARX modela sinhronog generatora se sa velikom tačnošću poklapaju sa rezultatima koji su dobijeni u Simulink-u.**

***Ključne reči*—nelinearni ARX model; sinhroni generator; sistem za regulaciju pobude.**

I. UVOD

ELEKTROENERGETSKI sistem je složen, dinamički sistem čija je glavna funkcija da sigurno, kvalitetno, pouzdano i ekonomično snabdijeva potrošače električnom energijom. Kompletan elektroenergetski sistem se sastoji iz podsistema proizvodnje, distribucije, prenosa i potrošnje električne energije. Proizvodnja električne energije se odvija u elektranama, u kojima se različiti oblici primarne energije, pomoću električnih generatora, transformišu u električnu energiju. Najčešće korišćeni električni generator je klasični sinhroni generator, dok se rjeđe u upotrebi mogu sresti asinhroni generator (najčešće dvostrano napajani), sinhroni generator sa stalnim magnetima, generator jednosmjerne struje itd. [1].

Brojne studije vezane za analizu elektroenergetskih sistema, kao što su studije stabilnosti, planiranja, testiranja sistema, analize dinamičkog odziva prilikom prelaznih procesa, zahtijevaju precizno i tačno modelovanje svake komponente sistema. Stoga, modelovanje sinhronog generatora, kao jednog od najvažnijih elemenata elektroenergetskog sistema, predstavlja veoma važan i zahtjevan zadatak [1], [2].

Zbog važnosti i velikog značaja modelovanja sinhronog generatora, brojne procedure za određivanje njegovih parametara su standardizovane i sublimirane u IEEE [3] i IEC standardima [4]. Osim toga, značajan broj naučnih radova je posvećen problematici modelovanja, odnosno estimacije parametara generatora. Metod baziran na naglom uklanjanju opterećenja je demonstriran u [5], dok je estimacija parametara na osnovu ogleda kratkog spoja prikazana u [6]. Procedura identifikacije parametara generatora kada se na ulaz dovode sinusoide različite frekvencije, tzv. *standstill frequency response test* (SSFR) demonstriran je u [7]. Jedan od modernijih i novijih identifikacionih metoda baziran je na podacima koji se dobijaju sa *phasor measurement unit*-a (PMU) (fazor napona, aktivna snaga, itd.) [8]. Takođe, u literaturi se mogu sresti i identifikacione procedure prilikom kojih se na namotaj pobude dovode signali različitih oblika: *chirp* signal [9], *sinc* signal [10], *pseudorandom binary sequence* (PRBS) signal [11], itd.

U ovom radu sinhroni generator se modeluje pomoću nelinearnog auto – regresionog modela sa spoljašnjim ulazom (ARX), čiji opis se može naći u [12], [13]. Kao ulazni signal u sinhroni generator koristi se napon pobude, dok je izlazni signal predstavljen preko napona na namotaju statora generatora. Ulazno – izlazni set podataka je dobijen modelovanjem sinhronog generatora, zajedno sa sistemom za automatsku regulaciju pobude generatora, u programskom paketu MATLAB Simulink. Na osnovu dobijenog seta podataka pomoću simulacionog modela, generator se modeluje pomoću pomenutog nelinearnog ARX modela, koji pokazuje izuzetno veliki stepen tačnosti. Sličan pristup se može sresti u [14], gdje se za modelovanje sinhronog generatora koristi takođe nelinearni ARX model, ali kod kojeg su nelinearnosti predstavljene stepenim funkcijama. Takođe, u pomenutom radu parametri modela su estimirani korišćenjem H∞ identifikacionog metoda.

Ovaj rad je organizovan na sljedeći način: u drugom poglavlju su date osnovne informacije o modelu koji je formiran u Simulink-u u cilju dobijanja seta ulazno-izlaznih mjerenja. Treće poglavlje je posvećeno prikazu korišćenog nelinearnog modela, dok su u četvrtom prikazani dobijeni rezultati. Na kraju, u Zaključku su ukratko sumirani rezultati i dati komentari vezani za budući rad.

Mihailo Micev – Elektrotehnički fakultet, Univerzitet Crne Gore, Bulevar Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (e-mail: mihailom@ ucg.ac.me).

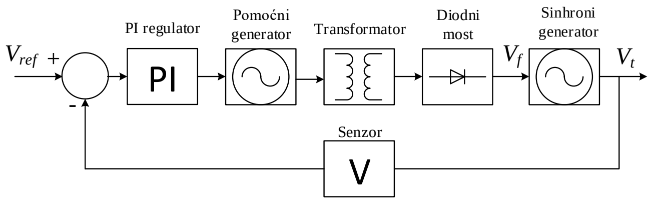
Martin Ćalasan – Elektrotehnički fakultet, Univerzitet Crne Gore, Bulevar Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (e-mail: martinc@ ucg.ac.me).

Milovan Radulović – Elektrotehnički fakultet, Univerzitet Crne Gore, Bulevar Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (e-mail: milovanr@ ucg.ac.me).

II. Simulacioni model sinhronog generatora

Kvalitet električne energije, kao i stabilnost elektroenergetskog sistema, u velikoj mjeri zavise od dva parametra, a to su napon i frekvencija. Frekvencija se smatra globalnim parametrom, jer je njena nazivna vrijednost jednaka u cijelom elektroenergetskom sistemu. Međutim, za napon se može reći da je lokalni parametar, jer njegova nazivna vrijednost zavisi od naponskog nivoa u svakom dijelu sistema pojedinačno. Zbog važnosti ova dva parametra, veoma je bitno održavati njihove vrijednosti na vrijednostima koje su bliske nazivnim, pri čemu je za napon dozvoljeno nešto veće odstupanje nego za frekvenciju. Regulacija frekvencije, tj. njeno održavanje na konstantnoj vrijednosti (50 ili 60 Hz, zavisno od sistema), vrši se pomoću turbinskog regulatora djelovanjem na mehaničku snagu vratila rotora. Odnosno, frekvencija primarno zavisi od tokova aktivnih snaga u elektroenergetskom sistemu, čime se formira takozvana *P-f* regulaciona kontura. Sa druge strane, djelovanjem na pobudu sinhronog generatora, može se uticati na tokove reaktivnih snaga, čime se definiše i naponski profil mreže. Na ovaj način može se definisati i druga regulaciona kontura koja predstavlja vezu između napona i reaktivnih snaga, odnosno *Q-U* kontura. Osim sinhronog generatora, za regulaciju tokova reaktivnih snaga mogu se koristiti redna i otočna baterija kondenzatora, sinhroni kompenzator, regulacioni transformator, statički VAr kompenzator, reaktor, itd.

U ovom radu se razmatra sinhroni generator kao najčešće korišćeno naponsko – reaktivno regulaciono sredstvo. Sinhroni generator je glavna komponenta sistema za regulaciju pobude sinhronog generatora, čija blok šema je prikazana na Slici 1.



Sl. 1. Blok šema sistema za regulaciju pobude sinhronog generatora.

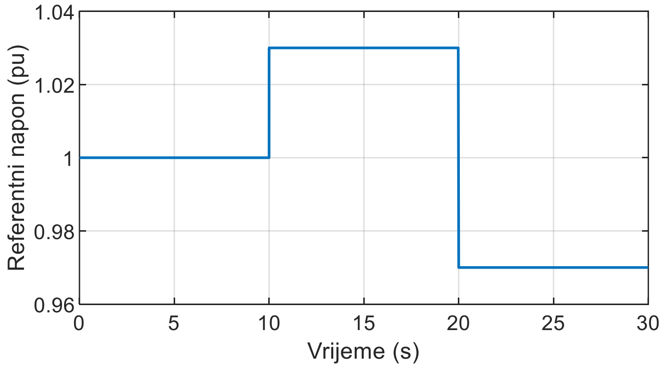
Prethodno prikazana šema je samo jedna od mogućih šema sistema za regulaciju pobude sinhronog generatora. Bitno je naglasiti da je prikazana struktura sistema za regulaciju pobude sinhronog generatora uprošćena u odnosu na onu koja se primjenjuje u stvarnim elektranama. Izlazni napon generatora V_t, tj. napon na namotajima statora, mjeri se pomoću senzora i upoređuje sa referentnim naponom V_{ref}. U komparatoru se, na osnovu te razlike, formira signal greške koji predstavlja ulaz u PI regulator. Dalje, regulator na svom izlazu daje upravljački signal, koji ustvari predstavlja pobudni napon za pomoćni sinhroni generator, čija snaga je značajno manja od snage glavnog sinhronog generatora. Pomoćni generator na svom izlazu daje naizmjenični napon, koji se pomoću transformatora i ispravljačkog diodnog mosta pretvara u jednosmjerni napon. Ovako dobijeni jednosmjerni napon V_f služi za napajanje pobudnog namotaja glavnog sinhronog generatora, tako da napon na izlazu što bolje prati referentni napon.

Prethodno opisana blok šema implementirana je u programskom paketu MATLAB Simulink. Podaci o korišćenim elementima su sljedeći:

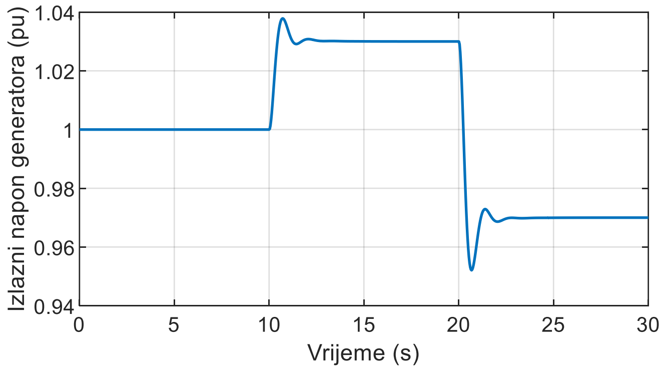
- Glavni sinhroni generator: S_n=2 MVA, U_n=400 V,
- Pomoćni sinhroni generator: S_n=8.1 kVA, U_n=400 V,
- Trofazni transformator: sprega Yd1, S_n=10 kVA, U_{n1}/U_{n2}=400 V/12 V,
- Diskretni PI regulator: K_p=10, K_i=8.
- Opterećenje koje napaja generator je čisto aktivno, snage 500 kW.

U sprovedenim simulacijama referentni napon V_{ref} je najprije podešen na nazivnu vrijednost (tj. 1 pu), zatim se nakon 10 s poveća na 1.03 pu, da bi se nakon dodatnih 10 s referentni napon smanjio na 0.97 pu. Tokom ovakvog režima rada generatora, neophodno je snimiti talasne oblike napona pobude V_f i izlaznog napona V_t. Cilj ovog rada je da se, na osnovu ovako snimljenih talasnih oblika, estimiraju parametri nelinearnog matematičkog modela generatora koji će biti opisan u narednom poglavlju.

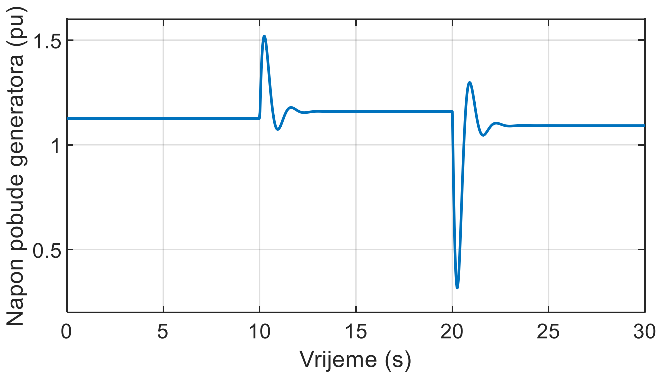
Kao što je i prethodno opisano, referentni napon generatora prikazan je na Slici 2. Nakon toga, na Slici 3 prikazan je izlazni napon generatora, dok je na Slici 4 prikazan napon pobude generatora u ovakvom režimu rada.



Sl. 2. Referentna vrijednost napona generatora.



Sl. 3. Izlazni napon generatora dobijen u Simulink-u.



Sl. 4. Napon pobude generatora dobijen u Simulink-u.

III. NELINEARNI ARX MODEL

U ovom poglavlju prikazan je nealinearni ARX model koji je korišćen u radu u cilju modelovanja sinhronog generatora na osnovu mjerenih ulaznih i izlaznih podataka. Set ulazno – izlaznih podataka neophodan za estimaciju parametara nelinearnog ARX modela dobijen je iz simulacije rada sinhronog generatora u Simulink-u, kao što je i opisano u prethodnom poglavlju.

U cilju boljeg razumijevanja nelinearnog modela, korisno je prvo ukratko prikazati linearni ARX model. Naime, ukoliko se sa $u(k)$ i $y(k)$ označe diskretna ulazna i izlazna mjerenja u trneutku k, respektivno, veza između izlaza i ulaza data je na sljedeći način:

$$A(z)y(k)=B(z)u(k)+e(k), \tag{1}$$

gdje $e(k)$ predstavlja poremećaj, a $A(z)$ i $B(z)$ su polinomi reda n_a i n_b, respektivno, po promjenljivoj z^{-1} koja predstavlja jedinično kašnjenje:

$$A(z)=1+a_1z^{-1}+a_2z^{-2}+\ldots+a_{n_a}z^{-n_a}=1+\sum_{i=1}^{n_a}a_iz^{-i}, \tag{2}$$

$$B(z)=b_1z^{-1}+b_2z^{-2}+\ldots+b_{n_b}z^{-n_b}=\sum_{i=1}^{n_b}b_iz^{-i}. \tag{3}$$

Koristeći prethodne definicije polinoma $A(z)$ i $B(z)$, (1) se može zapisati i na sljedeći način:

$$\left[1+\sum_{i=1}^{n_a}a_iz^{-i}\right]y(k)=\sum_{i=1}^{n_b}b_iz^{-i}u(k)+e(k), \tag{4}$$

odakle slijedi:

$$y(k)=-\sum_{i=1}^{n_a}a_iz^{-i}y(k)+\sum_{i=1}^{n_b}b_iz^{-i}u(k)+e(k), \tag{5}$$

ili uzimajući u obzir značenje operatora kašnjenja z^{-1}:

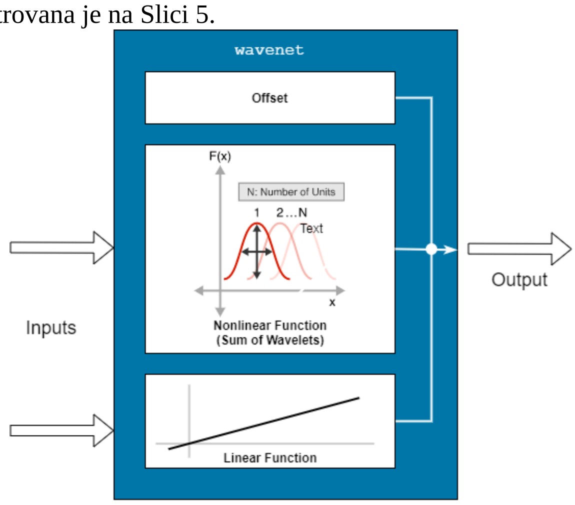
$$y(k)=-\sum_{i=1}^{n_a}a_iy(k-i)+\sum_{i=1}^{n_b}b_iu(k-i)+e(k). \tag{6}$$

Kao što se može zaključiti na osnovu (6), kod linearnog ARX modela izlaz $y(k)$ predstavlja linearnu kombinaciju izlaza u prethodnim trenucima $y(k-1)$, $y(k-2)$, ..., $y(k-n_a)$, kao i ulaza $u(k-1)$, $u(k-2)$, ..., $u(k-n_b)$. Ovi članovi se jednim imenom nazivaju regresori, tako da se može reći da je kod ARX modela izlaz u trenutku k linearna funkcija ostalih regresora.

Za razliku od prethodno opisanog modela, kod nelinearnog ARX modela postoji nelinearna funkcija mapiranja F između izlaza $y(k)$ i ostalih regresora:

$$y(k)=F\begin{bmatrix}y(k-1),y(k-2),\ldots,y(k-n_a),\\u(k-1),u(k-2),\ldots,u(k-n_b)\end{bmatrix}. \tag{7}$$

U ovom radu se kao nelinearna funkcija mapiranja F koristi tzv. *wavenet* funkcija, koja ustvari predstavlja mrežu, tj. sumu *wavelet* funkcija. Struktura ovakvog nelinearnog ARX modela ilustrovana je na Slici 5.



Sl. 5. Struktura nelinearnog ARX modela sa wavenet funkcijom mapiranja.

Sa matematičkog aspekta, *wavenet* mapiranje se može prikazati sljedećom relacijom:

$$y(k)=y_0+\left(X(k)-\bar{X}\right)^T PL+W\left[X(k)\right]+S\left[X(k)\right], \tag{8}$$

gdje $X(k)$ predstavlja vektor od ukupno m regresora, a $\bar{X}$ njegovu srednju vrijednost, y_0 je izlazni ofset (skalar), P je projekciona matrica dimenzija mxp (p – broj linearnih regresora), a L je vektor težinskih koeficijenata dimenzija $px1$. Nelinearni dio *wavenet* funkcije mapiranja ogleda se u funkcijama $W[X(k)]$ i $S[X(k)]$. Preciznije, funkcija $W[X(k)]$ predstavlja sumu proširenih i transliranih *wavelet*-a, dok funkcija $S[X(k)]$ predstavlja sumu proširenih i transliranih funkcija skaliranja, odnosno *scalelet*-a. Matematičke formulacije funkcija $W[X(k)]$ i $S[X(k)]$ definisane su pomoću (9) i (10):

$$W\left[X(k)\right]=\sum_{i=1}^{d_w}w_i f_w\left(b_i\left(X-\bar{X}\right)^T Q-c_i\right), \tag{9}$$

$$S\left[X(k)\right]=\sum_{i=1}^{d_s}s_i f_s\left(d_i\left(X-\bar{X}\right)^T Q-e_i\right). \tag{10}$$

Oznake koje se pojavljuju u (9) i (10) imaju sljedeća značenja:

- Q je projekciona matrica dimenzija mxq (q – proizvoljan parametar manji ili jednak od m),
- w_1 – w_{dw} su skalarni *wavelet* koeficijenti,
- s_1 – s_{ds} su skalarni *scaling* koeficijenti,
- b_1 – b_{dw} su skalari pod nazivom *wavelet* proširenja, dok su d_1 – d_{ds} skalari koji se zovu *scaling* proširenja koji množe ulaznu matricu,

- $c_1 - c_{dw}$ je vektor pod nazivom *wavelet* translacija,
- $e_1 - e_{ds}$ je vektor *scaling* translacija,
- f_w i f_s su funkcije definisane pomoću sljedećih relacija:

$$f_w(x) = e^{-xx^T/2}, \tag{11}$$

$$f_s(x) = \left(\dim(x) - xx^T\right)e^{-xx^T/2}. \tag{12}$$

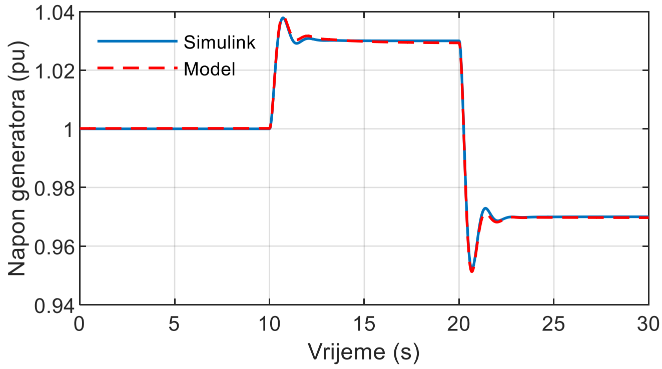
Estimacija parametara nelinearnog ARX modela izvršena je korišćenjem Levenberg – Marquardt algoritma [15], [16]. Ovaj algoritam je izuzetno pogodan za rješavanje nelinearnog problema najmanjih kvadrata prilikom tzv. *curve fitting*.

IV. Rezultati simulacija

U ovom poglavlju prikazanu su rezultati dobijeni modelovanjem sinhronog generatora pomoću nelinearnog ARX modela. Naime, blok šema opisana u drugom poglavlju implementirana je u Simulink-u i dobijeni su napon pobude i napon na izlazu generatora koji su prikazani na Slikama 3 i 4. Ovako dobijeni napon pobude i izlazni napon generatora formiraju set ulazno – izlaznih podataka pomoću kojeg se vrši estimacija parametara nelinearnog ARX modela.

Važno je ukazati da relacija (8) predstavlja najopštiju formu *wavenet* mapiranja kod nelinearnog ARX modela. U ovom radu nelinearni ARX model predstavljen je samo funkcijom $W[X(k)]$ iz relacije (8). Stoga, parametri koji su estimirani su matrica Q i vektori skalarnih *wavelet* koeficijenata w, wavelet proširenja b i wavelet translacija c. Korišćeno je m=28 regresora (n_a=14 i n_b=14), pri čemu su svi usvojeni kao nelinearni, a broj q je podešen na 19, čime je ispunjen uslov da je $q \leq m$.

Poređenje izlaznog napona generatora dobijenog u Simulink-u sa izlaznim naponom generatora koji je dobijen pomoću nelinearnog ARX modela dato je na Slici 6.

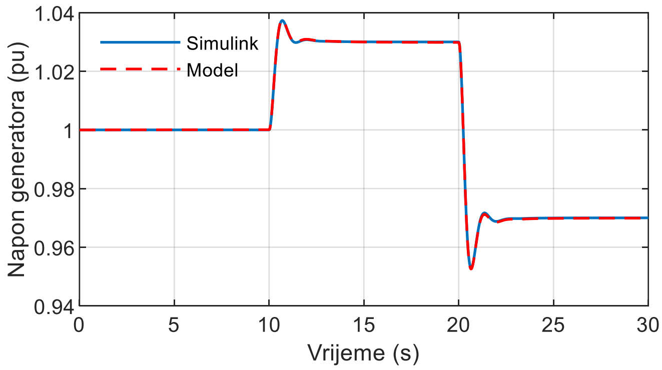


Sl. 6. Poređenje izlaznog napona generatora iz Simulink-a i izlaznog napona dobijenog primjenom nelinearnog ARX modela.

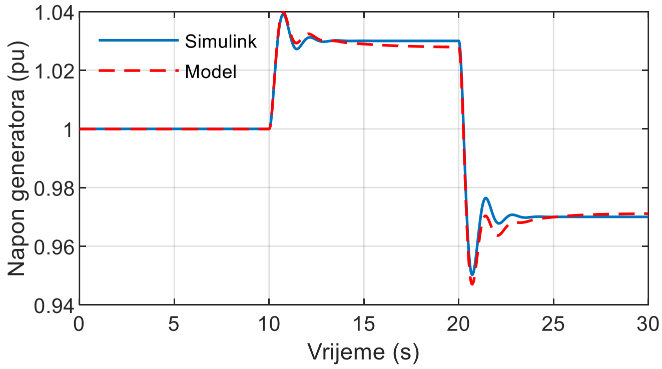
Na osnovu prethodno prikazanog grafika, jasno je da je korišćeni nelinearni ARX model izuzetno tačan i precizan jer su odstupanja od rezultata dobijenih pomoću Simulink-a veoma mala.

Nakon toga, ispitana je i mogućnost primjene nelinearnog ARX modela pri različitim uslovima u elektroenergetskom sistemu. Naime, estimacije parametara nelinearnog ARX modela izvršena je i u slučaju promjene opterećenja generatora. Razmatrana su dva slučaja: u prvom slučaju je opterećenje prepolovljeno, tj. smanjeno sa 500 kW na 250 kW, dok je u drugom slučaju opterećenje generatora udvostručeno, tj. iznosi 1000 kW. U oba slučaja izvršen je isti test: referentni napon je u početku jednak 1 pu, nakon 10 s se poveća na 1.03 pu, i nakon dodatnih 10 s se smanji na 0.97 pu. Odgovarajuća poređenja izlaznog napona iz Simulink-a sa naponom koji je dobijen primjenom nelinearnog ARX modela data su na Slikama 7 i 8. Na Slici 7 razmatran je slučaj opterećenja od 250 kW, dok je na Slici 8 analiziran slučaj kada je opterećenje generatora 1000 kW.

Kao što se može uočiti sa grafika, nelinearni ARX model omogućava odlično poklapanje rezultata sa onima dobijenim u Simulink-u i za različite uslove rada u elektroenergetskom sistemu. Time je pokazano da se ovaj model uspješno može adaptirati na promjenu uslova i da je pogodan za *online* primjenu u modelovanju sinhronog generatora.



Sl. 7. Poređenje izlaznog napona generatora iz Simulink-a i izlaznog napona dobijenog primjenom nelinearnog ARX modela (za opterećenje 250 kW).



Sl. 8. Poređenje izlaznog napona generatora iz Simulink-a i izlaznog napona dobijenog primjenom nelinearnog ARX modela (za opterećenje 1000 kW).

Na ovaj način pokazano je da estimirani parametri pri određenim uslovima rada (opterećenju na krajevima generatora) daju zadovoljavajuće rezultate i kada dođe do promjene radnih uslova, odnosno promjene opterećenja. Takođe, sa prethodnih grafika uočljivo je da je stepen poklapanja rezultata tokom dinamičkih procesa u prelaznom režimu veoma veliki.

V. Zaključak

U ovom radu razvijen je nelinearni ARX model pomoću kojeg se adekvatno može modelovati sinhroni generator.

Parametri ovog matematičkog modela estimirani su na osnovu napona pobude i napona na izlazu generatora, koji su dobijeni simulacijom sistema za regulaciju pobude sinhronog generatora u Simulink-u. Dobijeni rezultati pokazuju veliki stepen tačnosti i preciznosti korišćenog nelinearnog modela. U budućem radu biće razmotrena robusnost ovog modela u slučaju promjene parametara PI regulatora. Još jedan pravac istraživanja je primjena istog modela na eksperimentalno dobijene rezultate iz realne hidroelektrane.

LITERATURA

[1] T. A. Lipo, *Analysis of synchronous machines: Second Edition*. 2017.

[2] B. Adkins and R.G. Harley, The General Theory of Alternating Current Machines: Applications to Practical Problems, Springer-Science+Business Media, B.V.,1975.

[3] *IEEE Guide: Test Procedures for Synchronous Machines*, IEEE Standard 115-1995, 1995.

[4] *Rotating Electrical Machines—Part 4: Methods for Determining Synchronous Machine Quantities from Tests*, IEC Standard, IEC-60034-4 2008, 2008.

[5] M. Giesbrecht and L. A. E. Meneses, “Detailed derivation and experimental validation of a method for obtaining load conditions for salient pole synchronous machine quadrature axis parameters determination,” *IEEE Trans. Ind. Electron.*, vol. 66, no. 7, pp. 5049–5056, 2019.

[6] V. A. D. Faria, J. V. Bernardes, and E. C. Bortoni, “Parameter estimation of synchronous machines considering field voltage variation during the sudden short-circuit test,” *Int. J. Electr. Power Energy Syst.*, vol. 114, no. July 2019, p. 105421, 2020.

[7] T. L. Vandoorn, F. M. De Belie, T. J. Vyncke, J. A. Melkebeek, and P. Lataire, “Generation of multisinusoidal test signals for the identification of synchronous-machine parameters by using a voltage-source inverter,” *IEEE Trans. Ind. Electron.*, vol. 57, no. 1, pp. 430–439, 2010.

[8] B. Ahmadzadeh-Shooshtari, R. Torkzadeh, M. Kordi, H. Marzooghi, and F. Eghtedarnia, “SG parameters estimation based on synchrophasor data,” *IET Gener. Transm. Distrib.*, vol. 12, no. 12, pp. 2958–2967, 2018.

[9] M. Cisneros-González, C. Hernandez, R. Morales-Caporal, E. Bonilla-Huerta, and M. A. Arjona, “Parameter estimation of a synchronous-generator two-axis model based on the standstill chirp test,” *IEEE Trans. Energy Convers.*, vol. 28, no. 1, pp. 44–51, 2013.

[10] M. A. Arjona, M. Cisneros-Gonzalez, and C. Hernandez, “Parameter estimation of a synchronous generator using a sine cardinal perturbation and mixed stochastic deterministic algorithms,” *IEEE Trans. Ind. Electron.*, vol. 58, no. 2, pp. 486–493, 2011.

[11] M. Hasni, O. Touhami, R. Ibtiouen, M. Fadel, and S. Caux, “Estimation of synchronous machine parameters by standstill tests,” *Math. Comput. Simul.*, vol. 81, no. 2, pp. 277–289, 2010, doi: 10.1016/j.matcom.2010.05.010.

[12] https://www.mathworks.com/help/ident/index.html?s_tid=CRUX_lftnav

[13] B. Kovačević and G. Kvaščev, Identifikacija procesa, Akademska misao, Beograd, 2018.

[14] M. Dehghani, M. Karrari, W. Rosehart, and O. P. Malik, “Synchronous machine model parameters estimation by a time-domain identification method,” *Int. J. Electr. Power Energy Syst.*, vol. 32, no. 5, pp. 524–529, 2010, doi: 10.1016/j.ijepes.2009.07.010.

[15] K. Levenberg, “A Method for the Solution of Certain Non-Linear Problems in Least Squares,” *The Quarterly of Applied Mathematics*, vol. 2, pp 164-168, 1944.

[16] D.W. Marquardt, “An algorithm for least-squares estimation of nonlinear parameters,” *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431-441, 1963.

ABSTRACT

In this paper, the application of a nonlinear auto – regression model with external input (ARX) in order to model a loaded synchronous generator is presented. The input signal to the synchronous generator is represented as the field voltage, while the output signal is actually the terminal voltage of the generator. Therefore, to estimate the parameters of the adopted model, it is necessary to use the field voltage and the terminal voltage signals, which form the input – output data set. Additionally, the proposed experiment is based on changing the reference voltage of the synchronous generator excitation control system. The described procedure is implemented in the MATLAB Simulink software package. The results obtained using the proposed model of the synchronous generator coincide with high accuracy with the results obtained in Simulink.

Modelling of synchronous generator excitation control system using nonlinear ARX model

Mihailo Micev, Martin Ćalasan, Milovan Radulović

Identifikacija parametara mašine jednosmerne struje sa nezavisnom pobudom posle remonta

Miroslav Bjekić, Vojislav Vujičić, Marko Rosić, Marko Šućurović

***Apstrakt*— U radu je opisan postupak identifikacije parametara mašine jednosmerne struje sa nezavisnom pobudom posle izvršenog remonta. Snimljene su karakteristike praznog hoda generatora i mehaničke karakteristike motora za različite vrednosti priključnog napona, pobudne struje i dodate otpornosti u kolo indukta. Takođe su određeni i gubici usled obrtanja. Eksperimentalni rezultati potvrđeni su na dva načina koristeći merač momenta i elektromagnetnu kočnicu. Dobijeni koeficijenti elektromotorne sile (EMS) i momenta su upoređeni sa ranijim nazivnim vrednostima i dato je objašnjenje dobijenih razlika.**

***Ključne reči*— Mašina jednosmerne struje, identifikacija parametara, snimanje karakteristika**

I. UVOD

U laboratoriji za Električne mašine, pogone i automatiku - EMPA [1] na FTN u Čačku studenti IV godine modula Industrijska elektroenergetika studijskog programa Elektrotehničko i računarsko inženjerstvo izvode set vežbi iz predmeta Ispitivanje električnih mašina [2]. Programom predmeta se vrši ispitivanje mašine jednosmerne struje (JSS) i asinhrone mašine. Generacija studenata školske 2020/21 je dobila zadatak da ispita mašinu JSS posle remonta i da odredi njene nove parametre. Što tačnije određivanje parametara mašine dalje određuje kvalitet strujne i brzinske petlje odnosno dinamiku regulacije pogona sa mašinom JSS [3-5].

Naznačeni podaci ispitivane mašine JSS su sledeći:

U = 21,5 - 260 V
I_a = 17 - 17,4 A
n = 50 - 2741 o/min
P = 0,07 – 3,9 kW
J_{np} = 0,6 A, U_{np}=200 V
k_{TH} = 25 $(\text{Vmin})^{-1}$

Na osnovu navedenih naznačenih vrednosti mogu se odrediti dva bitna parametra mašine JSS sa nezavisnom pobudom: konstanta momenta – k_M i konstanta EMS – k_e. Na osnovu izraza za moment (1) i naznačenih podataka mašine JSS

$$M = c\Phi I_a = k_M I_a \qquad M = 9{,}55 P/n \tag{1}$$

može se dobiti vrednost konstante momenta koji iznose:

$$k_{M,\,P_{max}} = 9{,}55P/(I_a n) = 9{,}55\cdot 3900/(17{,}7\cdot 2741) = 0{,}7676\,\text{Nm/A}$$
$$k_{M,\,P_{min}} = 9{,}55P/(I_a n) = 9{,}55\cdot 70/(17\cdot 50) = 0{,}786\,\text{Nm/A}$$
$$\boldsymbol{k_M \approx 0{,}78\,\text{Nm/A}}.$$

Iz izraza za EMS:

$$E = U - R_a I_a \;\text{ gde je }\; E = c\Phi\omega = k_e n \tag{2}$$

dobija se vrednost konstante EMS

$$k_e = (U - R_a I_a)/n = (260 - 0{,}8\cdot 17{,}7)/2741 = 0{,}0897\,\text{Vmin}$$
$$\boldsymbol{k_e \approx 0{,}09\,\text{Vmin}}.$$

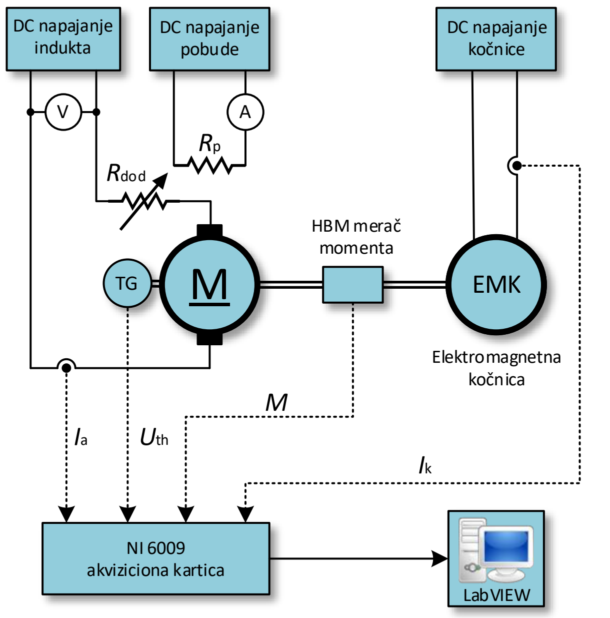
Može se primetiti da između izračunatim koeficijentima ne važi međusobna zavisnost da je k_M = 9,55 k_e, što je i bio dodatni razlog da se studentima postavi zadatak da se laboratorijskim ispitivanjem mašine jednosmerne struje tačno utvrde njegovi novi parametri.

Teorijske osnove studenti mogu da razumeju korišćenjem i simulacionih programa za njih posebno kreiranih u programskom paketu GEOGEBRA [6, 7]. Primeri koji su realizovani u laboratoriji i prikazani u nastavku rada su takođe realizovani u ovom programskom paketu [8].

II. KORIŠĆENA OPREMA I ELEKTRIČNA ŠEMA

Električna šema povezivanja motora JSS, elektromagnetne kočnice (EMK) i merno-akvizicione opreme je na slici 1. Na slici se može uočiti da je mašina JSS povezana na dva nezavisna izvora napajanja JSS, tj. realizovana je nezavisna pobuda pri izvođenju eksperimenta. Struja pobude je merena pomoću ampermetra. Napon napajanja indukta meren je voltmetrom, dok se struja indukta I_a merena pomoću strujnog senzora i akvizicione kartice NI 6009. Za potrebe eksperimenta, tj. za snimanje jedne od karakteristika mašine u kolu indukta redno je povezivana i promenljiva otpornost R_{dod}. Mašina JSS je mahanički povezana preko merača momenta sa EMK kojom je mašina opterećivana. Sa merača momenta se dobija vrednost momenta M čiji se naponski signal meri akvizicionom karticom. Za napajanje EMK korišćen je promenljivi izvor JSS čija struja I_k je merena pomoću šant otpornika i akvizicione kartice. Tako izmerena struja korišćena je za dobijanje vrednosti kočnog momenta EMK. Na drugoj strani vratila mašine se nalazi tahogenerator (TG) pomoću kod se tokom eksperimenta meri brzina obrtanja mašine. Naponski signal tahogeneratora U_{th} se takođe meri pomoću akvizicione kartice. Za potrebe merenja i snimanja navedenih veličina napravljena je odgovarajuća aplikacija u programu LabVIEW.

Miroslav Bjekić – Fakultet tehničkih nauka u Čačku, Univerzitet u Kragujevcu, Svetog Save 65, 32000 Čačak, Srbija (e-mail: mbjekic@gmail.com).
Vojislav Vujičić – Fakultet tehničkih nauka u Čačku, Univerzitet u Kragujevcu, Svetog Save 65, 32000 Čačak, Srbija (e-mail: vojislav.vujicic@ftn.kg.ac.rs).
Marko Rosić – Fakultet tehničkih nauka u Čačku, Univerzitet u Kragujevcu, Svetog Save 65, 32000 Čačak, Srbija (e-mail: marko.rosic@ftn.kg.ac.rs).
Marko Šućurović – Fakultet tehničkih nauka u Čačku, Univerzitet u Kragujevcu, Svetog Save 65, 32000 Čačak, Srbija (e-mail: marko.sucurovic@ftn.kg.ac.rs).

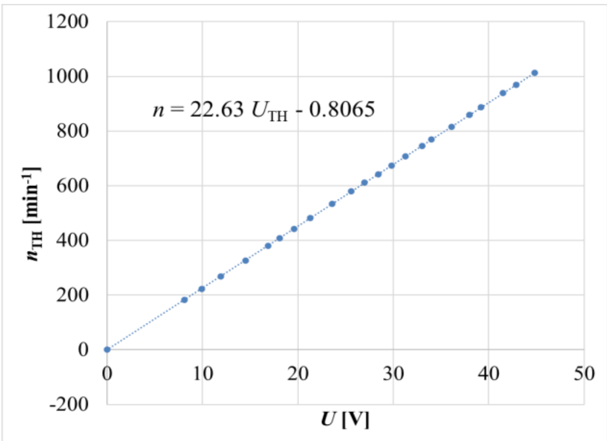


Sl. 1. Električna šema veze prilikom snimanja mehaničke karakteristike motora JSS sa nezavisnom pobudom

III. Karakteristike generatora JSS

A. *Baždarenje tahogeneratora*

Za potrebe merenja brzine obrtanja mašine korišćen je tahogenerator. Izvršeno je merenje napona na krajevima tahogeneratora za različite brzine obrtanja kako bi se dobila tačna vrednost njegove konstante (slika 2).

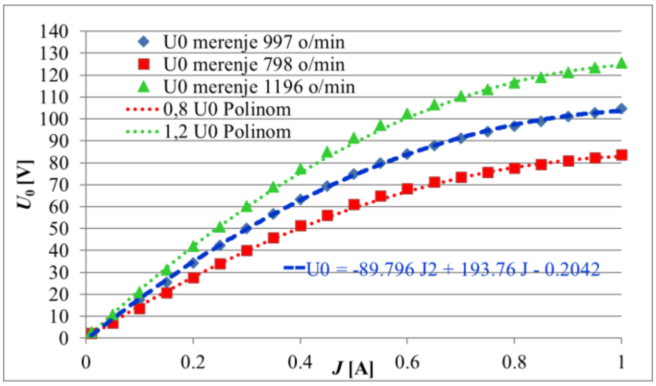


Sl. 2. Snimljena zavisnost brzine mašine JSS od napona na krajevima tahogeneratora

Dobijena je konstanta tahogeneratora od 22,63 (Vmin)$^{-1}$ koja je za 10,4% manja od naznačene vrednosti k_{TH}=25 (Vmin)$^{-1}$.

B. *Snimanje karakteristika praznog hoda*

Karakteristika praznog hoda je snimana stavljajući mašinu JSS u generatorski režim rada pokretanjem pomoću asinhrone mašine. Vršeno je snimanje karakteristike za tri brzine obrtanja: 997, 798 i 1196 o/min. Izvedena je interpolaciona funkcija zavisnosti napona praznog hoda od pobudne struje U_0= f(J) za brzinu 997 o/min. Znajući tu analitičku zavisnost moguće je pomoću proporcije dobiti je za bilo koju drugu brzinu. U ovom slučaju za brzine iznad i ispod osnovne krive polinom je množen je sa koeficijentom 0,8 i 1,2. Na slici 3 prikazane su izmerene vrednosti i interpolacione funkcije.



Sl. 3. Karakteristike praznog hoda generatora JSS sa nezavisnom pobudom i interpolirane funkcije za tri različite vrednosti pobudne struje

Na slici 3 se mogu uočiti odlična poklapanja krivih pa se zaključuje da se, sa dovoljnom preciznošću može usvojiti zavisnost napona praznog hoda (indukovane EMS) u funkciji pobudne stuje za konstantnu brzinu obrtanja kao:

$$U_0 = \left(-89,796J^2 + 193,76J - 0,2042\right)n/997 \qquad (3)$$

odakle sledi da konstante k_e=U_0/n za pobudne struje 0,6; 0,5 i 0,4 A iznose:

$$k_{e,\,J=0,6A} = \left(-89,796\cdot 0,6^2 + 193,76\cdot 0,6 - 0,2042\right)/997 = \underline{0,084\,\text{Vmin}}$$

$$k_{e,\,J=0,5A} = \left(-89,796\cdot 0,5^2 + 193,76\cdot 0,5 - 0,2042\right)/997 = 0,075\,\text{Vmin}$$

$$k_{e,\,J=0,4A} = \left(-89,796\cdot 0,4^2 + 193,76\cdot 0,4 - 0,2042\right)/997 = 0,063\,\text{Vmin}$$

Kao što je i očekivano, dobijene su tri različite vrednosti konstante EMS koja proporcionalno zavisi od jačine magnetnog fluksa koji stvara pobudni namotaj

$$E_a = L_p\, p'N'/60a'J\,n = k_e n \qquad (4)$$

gde je: L_p – induktivnost pobudnog navoja, p' – broj polova, N' – broj aktivnih provodnika navoja indukta, a' – broj paralelnih grana, J – pobudna struja i n – brzina obrtanja.

Na osnovu dobijenih koeficijenata i jednačine (4) se može odrediti veličina koja je srazmerna sa induktivnošću pobudnog namotaja C·L_p:

$$k_e/J = L_p p'N'/60a' = C\,L_p \;\Rightarrow\; C = p'N'/60a'. \qquad (5)$$

Vrednosti veličine C·L_p za tri različite vrednosti pobudne struje su:

$$C\,L_{p,\,J=0,6A} = k_{e,\,J=0,6A}/J = 0,084/0,6 = 0,14$$

$$C\,L_{p,\,J=0,5A} = k_{e,\,J=0,5A}/J = 0,075/0,5 = 0,15$$

$$C\,L_{p,\,J=0,4A} = k_{e,\,J=0,4A}/J = 0,063/0,4 = 0,1575$$

Dobijene vrednosti ukazuju na činjenicu da se sa povećanjem pobude ne dobija srazmerna vrednost magnetnog fluksa i da mašina ima nelinearnu karakteristiku magnećenja.

IV. Karakteristike motora JSS

A. *Određivanje konstante EMS motora JSS – k_e*

Prilikom snimanja mehaničke karakteristike motora JSS pomoću akvizicione opreme merene su sledeće veličine: pobudna struja – J, struja opterećenja – I, napon tahogeneratora – U_{th}, moment merenog pomoću merača momenta – M_{mer} i moment dobijen na osnovu merene vrednosti struje EMK – M_k.

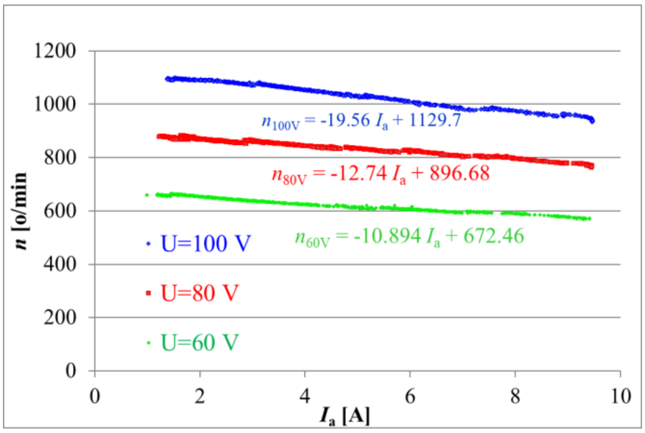
Izvršena su merenja pri promeni:

a) Napona napajanja: U=100, 80 i 60 V
b) Dodate otpornosti u kolu indukta: R_{dod}=0, 15, 27 Ω
c) Pobudne struje: J=0,6, 0,5 i 0,4 A

Svi rezultati su obrađeni i prikazani grafički u nastavku rada.

a) Promena napona, pobudna struja J=0,6 A

Izmerene vrednosti brzine u funkciji struje indukta, za tri različita napona i pri konstantnoj struji pobude date su na slici 4.



Sl. 4. Zavisnost napona motora od struje opterećenja pri različitim vrednostima priključnog napona

Analitički izraz zavisnosti brzine motora od struje opterećenja je:

$$n = \mathrm{f}(I_a) = n_0 - R_a I_a / k_e = U/k_e - R_a I_a / k_e \qquad (6)$$

odakle se na osnovu interpoliranih pravih čiji su izrazi prikazani na slici 4 dobija konstanta EMS:

$$k_{e,\ U=100V} = 100/1129,7 = 0,088519 \approx \underline{0,089\,\mathrm{Vmin}}$$
$$k_{e,\ U=80V} = 80/896,68 = 0,089218 \approx \underline{0,089\,\mathrm{Vmin}}$$
$$k_{e,\ U=60V} = 60/672,46 = 0,089224 \approx \underline{0,089\,\mathrm{Vmin}}$$

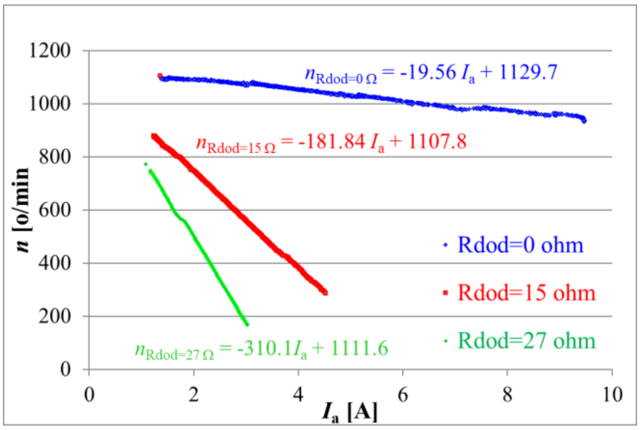
b) Dodavanje otpora u kolo indukta, U=100 V, J=0,6 A

Izmerene vrednosti brzine u funkciji struje indukta, za tri različite vrednosti dodatih otpora u kolo indukta, pri konstantnom naponu i struji pobude date su na slici 5.

Iz prikazanih interpolacionih funkcija (slika 5) može se odrediti konstanta EMS:

$$k_{e,\ Rdod=0\Omega} = 100/1129,7 = 0,08852 \approx \underline{0,089\,\mathrm{Vmin}}$$
$$k_{e,\ Rdod=15\Omega} = 100/1107,8 = 0,09027 \approx \underline{0,09\,\mathrm{Vmin}}$$
$$k_{e,\ Rdod=27\Omega} = 100/1111,6 = 0,08996 \approx \underline{0,09\,\mathrm{Vmin}}$$

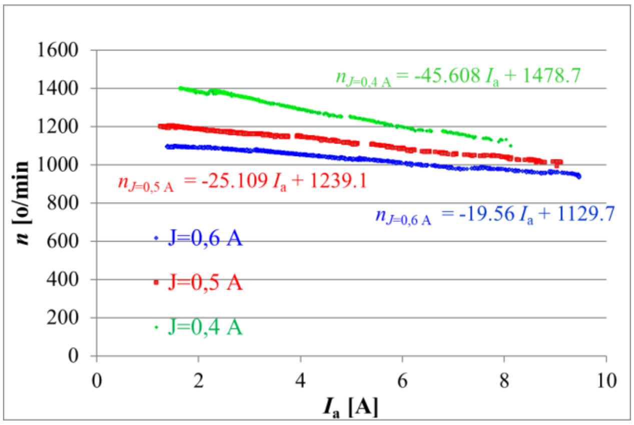
Dobijene vrednosti potvrđuju da konstanta EMS ne zavisi od dodate vrednosti otpora u kolo indukta. Ove tri podjednake vrednosti potvrđuju tačnost izvršenog merenja.



Sl. 5. Zavisnost napona motora od struje opterećenja pri različitim vrednostima dodate otpornosti u kolo indukta

c) Promena pobudne struje, U=100 V

Izmerene vrednosti brzine u funkciji struje indukta, za tri različite vrednosti struje pobude i pri konstantnom naponu date su na slici 6.



Sl. 6. Zavisnost napona motora od struje opterećenja pri različitim vrednostima pobudne struje

Na osnovu dobijenih interpolacionih pravih dobijaju se sledeće vrednosti konstante EMS:

$$k_{e,\ J=0,6A} = 100/1129,7 = 0,088519 \approx \underline{0,089\,\mathrm{Vmin}}$$
$$k_{e,\ J=0,5A} = 100/1239,1 = 0,08070 \approx 0,081\,\mathrm{Vmin}$$
$$k_{e,\ J=0,4A} = 100/1478,7 = 0,06763 \approx 0,068\,\mathrm{Vmin}$$

Iz izračunatih vrednosti može se primetiti da sa povećanjem pobudne struje se povećava i konstantna EMS motora pošto je direktno srazmerna sa magnetnim fluksom koji stvara pobuda. Ako se ova konstantna podeli sa pobudnom strujom dobija se vrednost srazmerna sa induktivnošću pobudnog namotaja:

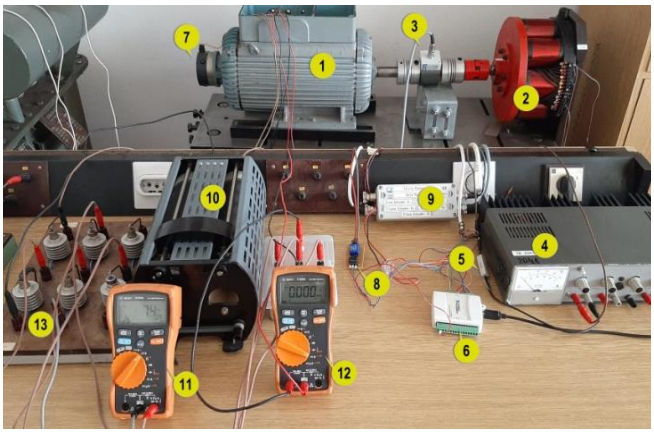
$$C\,L_{p,\ J=0,6A} = k_{e,\ J=0,6A} / J = 0,089/0,6 = 0,148$$
$$C\,L_{p,\ J=0,5A} = k_{e,\ J=0,5A} / J = 0,081/0,5 = 0,162$$
$$C\,L_{p,\ J=0,4A} = k_{e,\ J=0,4A} / J = 0,068/0,4 = 0,17$$

B. Određivanje konstante momenta motora JSS – k_M

Za određivanje konstante momenta jednosmernog motora neophodno je obezbediti što tačnije merenje momenta. Ono je izvršeno na dva načina:

- Pomoću merača momenta i
- Pomoću baždarene EMK [9].

Na slici 7 prikazana oprema korišćena za realizaciju snimanja mehaničkih karakteristika motora JSS. Motor JSS (1) je spregnut sa EMK (2) preko merača momenta HBM T22/50NM (3) pomoću kog se direktno meri moment. EMK se napaja sa izvora JSS (4) čija struja određuje kočni moment. Struja kočenja je merena pomoću šant otpornika od 10 Ω (5). Na akvizicionu karticu NI 6009 (6) povezani su merni (naponski) signali sa tahogeneratora (7) za merenje brzine, zatim sa senzora Ametes CS10A-02 (8) za merenje struje indukta, sa priključne kutije merača momenta (9) za vrednost momenta, i sa šant otpornika za merenje struje kočenja (5). Od ostale opreme, na slici 7 prikazani su: promenljivi otpornik dodavan u kolo indukta (10), voltmetar za merenje napona napajanja motora JSS (11), ampermetar (12) za merenje struje pobude i ispravljačko kolo (13) za napajanje kola indukta.



Sl. 7. Oprema korišćena za merenje mehaničke karakteristike motora JSS

Povezivanjem na akvizicionu karticu, izvršeno je podešavanje analognih naponskih ulaza u okviru softvera LabVIEW, nakon čega je kreirana aplikacija za snimanje signala.

Pored direktnog merenja elektromagnetnog momenta pomoću HBM merača momenta iskorišćena je i EMK koja je imala dvostruku funkciju:

- da se pomoću nje kontrolisano opterećuje motor JSS, i
- da se pomoću poznate karakteristike EMK dođe do vrednosti elektromagnetnog momenta motora znajući analitičku zavisnost kočnog momenta u funkciji brzine obrtanja diska i struje kočnice [9]:

$$M_k = f(n) = A \cdot e^{(B \cdot n)} + C \cdot e^{(D \cdot n)} \quad (7)$$

$$\begin{aligned} A &= A_1 \cdot I_k^3 + A_2 \cdot I_k^2 + A_3 \cdot I_k + A_4 \\ B &= B_1 \cdot I_k^3 + B_2 \cdot I_k^2 + B_3 \cdot I_k + B_4 \\ C &= C_1 \cdot I_k^3 + C_2 \cdot I_k^2 + C_3 \cdot I_k + C_4 \\ D &= D_1 \cdot I_k^3 + D_2 \cdot I_k^2 + D_3 \cdot I_k + D_4 \end{aligned} \quad (8)$$

Izračunate vrednosti koeficijenta polinoma napisane u matričnom obliku za određivanje kočnog momenta kočnice iznose [9]:

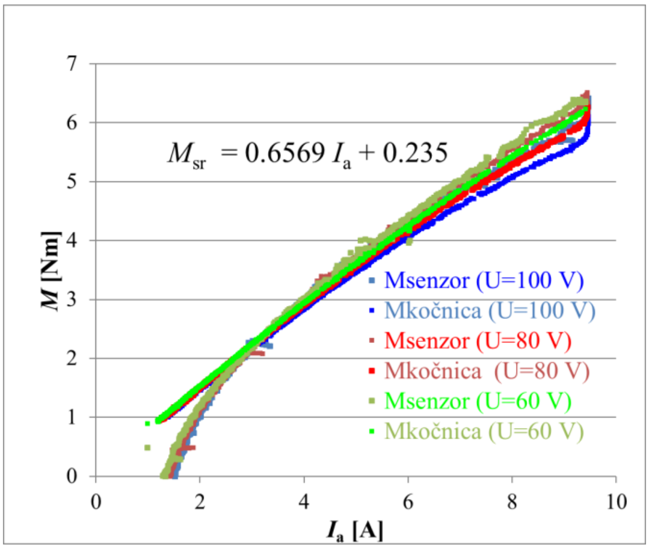
$$\begin{bmatrix} A_1 & B_1 & C_1 & D_1 \\ A_2 & B_2 & C_2 & D_2 \\ A_3 & B_3 & C_3 & D_3 \\ A_4 & B_4 & C_4 & D_4 \end{bmatrix} = \begin{bmatrix} -4,741 & -0,003162 & 3,528 & -0,03787 \\ 26,41 & 0,007258 & -24,95 & 0,01937 \\ 2,245 & -0,007258 & -2,747 & 0,004357 \\ -0,2562 & 0,001099 & 0,3187 & -0,0072 \end{bmatrix}$$

Pomoću datih koeficijenata dobijaju se mehaničke karakteristike za bilo koju struju kočenja <u>iz mernog opsega brzine obrtanja kočnice</u>.

Rezultati merenja elektromagnetnog momenta pomoću merača momenta (M_{senzor}) i posredno – preko poznatih parametara EMK ($M_{kočnica}$) su prikazani na slikama:

a) promeni napona, slika 8
b) dodavanje otpora u kolo indukta, slika 9, i
c) promeni pobudne struje, slika 10.

a) Promena napona, pobudna struja J=0,6 A



Sl. 8. Zavisnost momenta motora od struje opterećenja pri različitim vrednostima priključnog napona

Na osnovu grafički prikazanih vrednosti momenta (slika 8) za tri različita napona napajanja dobijena je srednja vrednost funkcije mehaničke karakteristike. Iz izraza za moment:

$$M = k_M I_a \Rightarrow k_M = M / I_a \quad (9)$$

Dobijena je srednja vrednost konstante EMS koja iznosi

$$k_{M,\,sr} = 0{,}6569.$$

Analizirajući prikazane vrednosti izmerenog momenta može se zaključiti sledeće:

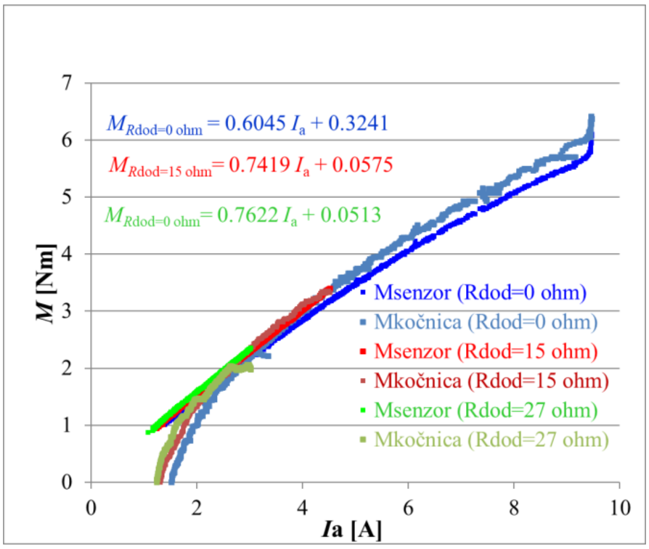
- Dobijeni rezultati pokazuju zadovoljavajuća poklapanja između rezultata merenja elektromagnetnog momenta pomoću merača momenta i posredno preko EMK. Time je još jednom potvrđena zadovoljavajuća tačnost posredno izračunate vrednosti momenta preko brzine obrtanja i struje kočnice;

- Vrednost promenljivog priključnog napona nije ni trebala da utiče na oblik dobijenih krivih, kao što i prikazane krive pokazuju;
- Vrednosti momenata dobijenih pomoću EMK za male vrednosti struja odstupaju od linearnog oblika, što se i očekivalo znajući, ranije utvrđen, merni opseg kočnice [9];
- Može se uočiti da krive ne polaze iz koordinatnog početka, već da imaju određenu vrednost momenta za struju indukta I_a=0. Ova vrednost ne može da se dobije merenjem, pošto u režimu praznog hoda, tj. neopterećene mašine, postoji određena struja indukta kojom se pokrivaju gubici usled obrtanja.

Sve su to razlozi zbog kojih postoji neslaganje između, na početku rada izračunate vrednosti konstante momenta motora i izmerenih vrednosti momenta. Zaključak je da se za motor JSS ne može primeniti zavisnost $M=k_mI$ i da naznačene vrednosti važe za naznačene radne režime.

Neslaganja između naznačenih i merenih vrednosti se mogu tumačiti i činjenicom da je mašina bila remontovana, pa se moglo očekivati i da su joj fabrički naznačene vrednosti promenjene.

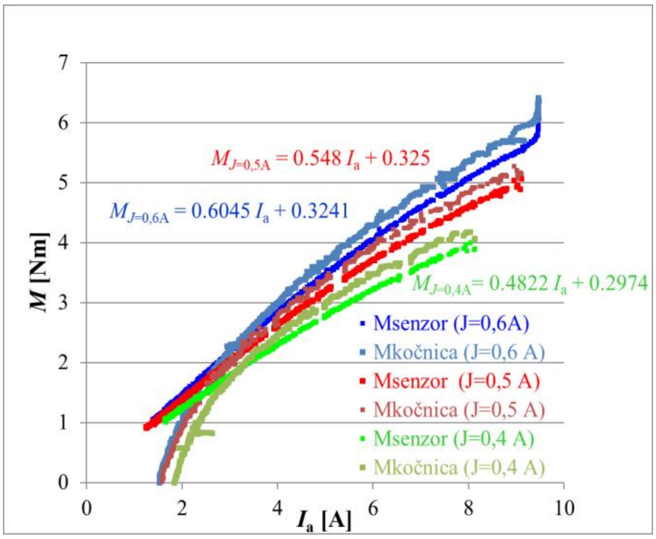
b) Dodavanjem otpora u kolo indukta, U=100 V, J=0,6 A



Sl. 9. Zavisnost momenta motora od struje opterećenja pri različitim vrednostima dodane otpornosti u kolu indukta

Dobijene vrednosti konstante momenta motora JSS za dodate vrednosti otpornosti u kolo indukta (slika 9) pokazuju da su koeficijenti za slučaj dodavane otpornosti u kolo indukta približne vrednosti k_M=0,75 Nm/A što se poklapa sa, na početku rada, izračunatom konstantom momenta na osnovu naznačenih vrednosti (0,78 Nm/A). Ovo poklapanje ukazuje da je potrebno posebno obratiti pažnju na otpornost kola indukta i izvršiti dodatna ispitivanja, što će biti tema budućeg rada.

c) Promena pobudne struje, U=100 V



Sl. 10. Zavisnost momenta motora od struje opterećenja pri različitim vrednostima pobudne struje

Sa grafika datih na slici 10 određene su vrednosti koeficijenta pravca mehaničkih karakteristika koje za pojedine pobudne struje iznose:

$$k_{M,\ J=0,6A} = 0,6045$$
$$k_{M,\ J=0,5A} = 0,548$$
$$k_{M,\ J=0,4A} = 0,4822$$

pomoću kojih se može odrediti i induktivnost pobude:

$$L_{p,\ J=0,6A} \sim k_{M,J=0,6A}/9,55J = 0,6045/9,55 \cdot 0,6 = 0,105$$
$$L_{p,\ J=0,5A} \sim k_{M,J=0,5A}/9,55J = 0,548/9,55 \cdot 0,5 = 0,116$$
$$L_{p,\ J=0,4A} \sim k_{M,J=0,4A}/9,55J = 0,4822/9,55 \cdot 0,4 = 0,126.$$

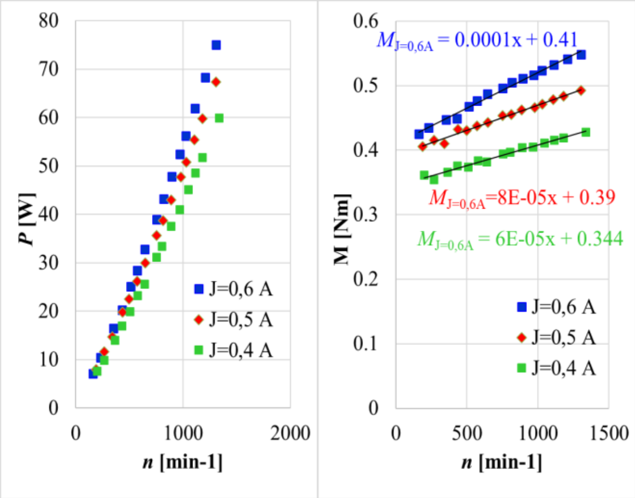
Ovako dobijene induktivnosti potvrđuju činjenicu da je magnetno kolo mašine JSS nelinearno i da sa povećanjem pobude se smanjuje induktivnost (mašina ulazi u zasićenje).

Pošto su dobijene procenjene vrednosti induktivnosti na tri različita načina: preko karakteristike praznog hoda, mehaničke karakteristike n=f(I_a) i n=f(M), sve vrednosti su zajedno prikazane u tabeli I.

TABELA I
INDUKTIVNOSTI POBUDNOG NAMOTAJA

C L_p [H]	J=0,6 A	J=0,5 A	J=0,4 A
K-ka U_0=f(J)	0,14	0,15	0,16
K-ka n=f(I_a)	0,15	0,16	0,17
K-ka M=f(I_a)	0,11	0,12	0,13

Na slici 11 prikazane su zavisnosti gubitaka i momenta usled obrtanja za tri različite vrednosti pobudne struje. Čime je potvrđeno da se zavisnost elektromagnetnog momenta ne može izraziti kao linearna funkcija struje indukta, već ima složeniji oblik.



Sl. 11. Snimljene karakteirstike gubitaka usled obrtanja i momenta usled obrtanja za različite vrednosti pobudne struje

V. Zaključak

U radu je izvšena identifikacija parametara mašine JSS posle izvšenog remonta. Najznačajniji parametar koji je trebalo eksperimentalnim putem dobiti i proveriti u odnosu na naznačene vrednosti je konstanta EMS i konstanta momenta k_M. Ona su izračunate obrađujući rezultate merenja u ogledu praznog hoda i ogledu snimanja mehaničke karakteristike motora JSS sa nezavisnom pobudom, za 3 načina prilagođenja karakteristike motora (promenom napona, slabljenjem polja i dodavanjem otpornosti.

Utvrđeno je da postoje neslaganja između, na početku rada izračunate vrednosti konstante EMS i konstante momenta motora jednosmerne struje preko naznačenih parametara i izmerenih vrednosti. U radu je potvrđeno da se moment motora ne može za ceo radni opseg predstaviti kao linearna zavisnost struje opterećenja ($M \neq k_M I_a$) i da se preciznija zavisnost dobija uzimanjem u obzir gubitaka usled obrtanja i momenta koji postoji i u režimu praznog hoda.

Takođe, u radu je potvrđena tačnost postupka određivanja momenta motora preko parametara elektromagnetne kočnice (struje i brzine) kojom se kontrolisano opterećuje.

Zahvalnica

Istraživanja prezentovana u ovom radu su delimično finansirana sredstvima Ministarstva prosvete, nauke I tehnološkog razvoja RS, ugovor br. 451-03-9/2021-14/200132 čiji je realizator Fakultet tehničkih nauka u Čačku – Univerziteta u Kragujevcu.

Literatura

[1] Sajt Laboratorije za električne mašine, pogone i automatiku – FTN u Čačku: www.empa.ftn.kg.ac.rs

[2] Sajt predmeta Ispitivanje električnih mašina – Laboratorija EMPA, FTN u Čačku: http://www.empa.ftn.kg.ac.rs/osnovne-studije/ispitivanje-elektricnih-masina.php

[3] M. Rosić, S. Antić, M. Bjekić, V. Vujičić, "Educational laboratory setup of DC motor cascade control based on dSPACE1104 platform", Proceedings of 5th International Conference on Renewable Electrical Power Sources - ICREPS2017, vol. 5 no. 1, pp. 213-222, DOI:10.24094/mkoiee.017.5.1.213, Belgrade, Serbia, 12-13 October, 2017.

[4] D. T. Liem and K. K. Ahn, "DC motor parameters identification and sensorless torque estimation using Fuzzy PID," 2012 12th International Conference on Control, Automation and Systems, 2012, pp. 76-81.

[5] S. S. Saab, R. Abi Kaed-Bey, "Parameter Identification of a DC Motor: An Experimental Approach", ICECS 2001. 8th IEEE International Conference on Electronics, Circuits and Systems. Malta, 2-5 Sept. 2001.

[6] Spisak autora i programa koji su kreirali softvere u GEOGEBRI iz oblasti Električne mašine i pogoni. Dostupno na: www.empa.ftn.kg.ac.rs/razno/GEOGEBRA_ELEKTRICNE_MASINE.html

[7] M. Bjekic, Kreirani programi u GEOGEBRI. Dostupno na: https://www.geogebra.org/u/mbjekic

[8] DC motor separate exited, https://www.geogebra.org/m/m2mwpeja M. Bjekić, M. Božić, M. Rosić, M. Šućurović, „Elektromagnetna kočnica sa jednim obrtnim diskom za laboratorijska ispitivanja električnih motora", Tehničko rešenje, Fakultet tehničkih nauka u Čačku, 2015. Dostupno na: http://www.ftn.kg.ac.rs/docs/resenja/EM_kocnica.pdf

Abstract

The paper describes identification process of separately excited DC machine parameters after the machine repair was performed. Mechanical and no load characteristics of DC machine are recorded with different values of armature voltages, excitation current and added resistance in the armature circuit. Rotational and ventilation losses of the machine are determined as well. Experimental results are confirmed in two ways: using torque sensor and using calibrated electromechanical brake. Machine EMF and torque coefficients are compared with previous rated machine values after which a proper discussion of obtained differences was given.

Separately excited DC machine parameters identification after the machine repair

M. Bjekic, V. Vujicic, M. Rosic, and M. Sucurovic

Simulacija histerezisnih petlji interpolacijom harmonijskih komponenti magnetskog polja

Srđan Divac, Branko Koprivica

***Apstrakt*—Cilj ovog rada je prikaz postupka simulacije histerezisnih petlji feromagnetskog uzorka koja je sprovedena simulacijom vremenskih oblika jačine magnetskog polja (u nastavku magnetskog polja) i magnetske indukcije. Podaci za magnetsko polje i indukciju su dobijeni merenjem, pri kontrolisanom obliku magnetske indukcije, za dva razmatrana oblika - sinusni i trougaoni. Na osnovu izmerenih podataka su određeni harmonici magnetskog polja za poznate amplitude magnetske indukcije. Nove vrednosti ovih harmonika se određuju interpolacijom prethodno izračunatih vrednosti, za amplitudu magnetske indukcije od interesa. Na osnovu ovih harmonika se vrši simulacija novog vremenskog oblika magnetskog polja. Nova histerezisna petlja se simulira korišćenjem ovog magnetskog polja i nove simulirane magnetske indukcije odgovarajućeg oblika. U radu je opisan navedeni postupak simulacije, dato je poređenje interpoliranih i izračunatih amplituda harmonika i vremenskih oblika magnetskog polja, kao i histerezisnih petlji. Takođe, u radu je data odgovarajuća analiza i diskusija rezultata.**

***Ključne reči* — Magnetsko polje; Histerezisna petlja; Harmonijska analiza; Interpolacija; Simulacija.**

I. Uvod

Modelovanje ili simulacija procesa magnećenja feromagnetskog uzorka zahteva veliki broj izmerenih podataka od značaja za dati proces [1-8], kako bi se odredili parametri modela (Preisach [1], Jiles-Atherton [2], Tellinen [3], Milovanović-Koprivica [4] i drugi) ili izvršilo treniranje neuronske mreže [6-7]. Proces magnećenja feromagnetskog materijala se može uspešno modelovati pomoću harmonijskih komponenti [9]. Određivanje parametara modela, treniranje neuronske mreže i drugi simulacioni postupci su neretko zahtevni u matematičkom i vremenskom pogledu. Dodatno, njihova tačnost i mogućnosti šire primene su ograničeni. Za prevazilaženje navedenih problema su potrebni alternativni postupci simulacije magnetskih veličina i histerezisa.

U ovom radu izložen je jedan alternativni postupak simulacije histerezisnih petlji baziran na interpolaciji amplituda i faza konačnog broja harmonika izmerenih magnetskih polja za magnetsku indukciju poznatog oblika i amplitude.

Podaci potrebni za proračun dobijeni su merenjem magnetskog polja i indukcije torusnog uzorka, napravljenog od orijentisanog feromagnetskog lima, primenom merne metode bazirane na personalnom računaru [10]. Pri merenjima su korišćena dva oblika magnetske indukcije - sinusni i trougaoni. U kontrolisanom merenju su postignute amplitude u opsegu od 0,2 T do 1,6 T sa korakom od 0,2 T, za oba razmatrana oblika. Merni podaci za verifikaciju rezultata proračuna dobijeni su pri istim oblicima magnetske indukcije za amplitude 0,5 T, 0,9 T i 1,5 T.

U radu su predstavljene izmerene histerezisne petlje za oba razmatrana oblika magnetske indukcije, kao i amplitude i faze harmonika odgovarajućih magnetskih polja. Takođe, prikazana su poređenja interpoliranih i izračunatih amplituda harmonika magnetskih polja, vremenskih oblika magnetskog polja i histerezisnih petlji za sve razmatrane slučajeve. Izvršena je analiza odstupanja amplituda harmonika i površina histerezisnih petlji i mogućnosti primene predloženog simulacionog postupka, uz odgovarajuću diskusiju.

II. Simulacioni postupak

Magnetsko polje $H(t)$ je fizička veličina koja se može predstaviti pomoću neparnih harmonijskih komponenti u obliku sume konačnog broja harmonika [11], i to:

$$H(t) = \sum_{i=1}^{N} \left(H_i \cos\left((2i-1)\omega t + \alpha_i \right) \right), \quad (1)$$

gde su H_i i α_i amplituda i faza i-tog harmonika, respektivno, i je red harmonika, N je najviši red, ω je kružna učestanost osnovnog harmonika i t je vreme.

Simulacija histerezisne petlje od značaja izvodi se kroz postupak sa sledećim koracima:

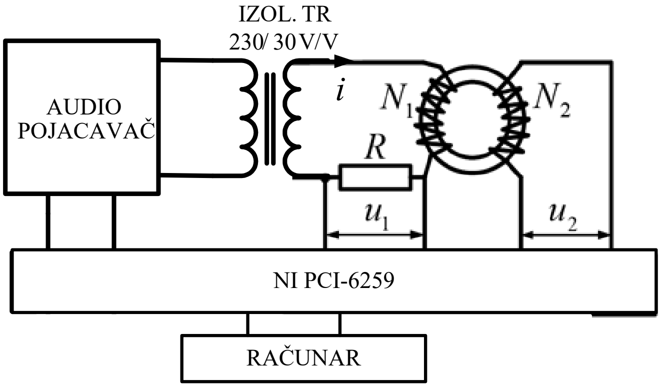
1. proračun N amplituda i faza neparnih harmonika izmerenih magnetskih polja za odgovarajući opseg amplituda magnetske indukcije,
2. interpolacija novih amplituda i faza harmonika polja za amplitude magnetske indukcije od interesa,
3. proračun vremenskog oblika magnetskog polja primenom (1) i simulacija vremenskog oblika magnetske indukcije i
4. simulacija histerezisnih petlji na osnovu rezultata iz prethodnog koraka.

III. Merna metoda i aparatura

Vremenski oblici magnetskog polja i indukcije dobijeni su primenom metode merenja bazirane na personalnom računaru i akviziciji podataka [10]. Merenja su sprovedena sa namotanim torusnim uzorkom od feromagnetskog lima oznake 27PH100, proizvođača POSCO. Blok šema merno-akvizicionog sistema je prikazana na Sl. 1.

Srđan Divac – Fakultet tehničkih nauka u Čačku, Univerzitet u Kragujevcu, Svetog Save 65, 32000 Čačak, Srbija, (e-mail: divacsrdjan@gmail.com) – Student doktorskih studija - Stipendista Ministarstva.

Branko Koprivica – Fakultet tehničkih nauka u Čačku, Univerzitet u Kragujevcu, Svetog Save 65, 32000 Čačak, Srbija, (e-mail: branko.koprivica@ftn.kg.ac.rs).



Sl. 1. Blok šema merno-akvizicionog sistema za merenje magnetskih karakteristika.

Merno-akvizicioni sistem se sastoji od naponski kontrolisanog izvora naizmeničnog napona, šant otpornika od R=0,5 Ω, torusnog uzorka, akvizicione kartice i računara.

Naponski kontrolisan izvor čini audio pojačavač CROWN XLi 2500 i izolacioni transformator 230/30 V/V, zajedno sa analognim izlazom akvizicione kartice NI PCI-6259. Ova kartica se koristi i za merenje napona u_1 na šant otporniku i indukovanog napona u_2 na krajevima indukcionog namotaja torusnog uzorka. Za potrebe merenja napravljena je aplikacija u programskom paketu LabVIEW koja vrši obradu i čuvanje izmerenih podataka. Takođe, ova aplikacija se koristi za kontrolu vremenskog oblika magnetske indukcije, određene na osnovu izraza (3) - integrala po vremenu izmerenog napona u_2, prema zadatom vremenskom obliku indukcije. Kontrola amplitude i oblika magnetske indukcije je sprovedena uspostavljanjem digitalne (programske) povratne sprege [12]. Kroz iterativni postupak kontrole se menja napon izvora sve dok se ne ostvari zadati vremenski oblik indukcije.

Merenje napona u_1 i u_2 je realizovano sa 1000 mernih podataka po periodi signala. Sve ostale veličine koje se računaju na osnovu ovih napona takođe sadrže ovaj broj podataka.

Struja $i(t)$ kroz pobudni namotaj meri se indirektno, kao količnik izmerenog napona šant otpornika u_1 i njegove otpornosti R. Ova struja se koristi za proračun trenutne vrednosti magnetskog polja prema Amperovom zakonu [13]:

$$H(t)=\frac{N_1 i(t)}{l}\quad\left[\frac{\mathrm{A}}{\mathrm{m}}\right], \qquad (2)$$

gde je N_1 broj navojaka pobudnog namotaja, a l dužina srednje linije jezgra.

Napon u_2 koji se indukuje na indukcionom namotaju torusnog uzorka proporcionalan je brzini promene ukupnog fluksa u jezgru, a magnetska indukcija se može proračunati pomoću izraza:

$$B(t)=\frac{1}{N_2 A}\int_0^t u_2(\tau)\,\mathrm{d}\tau \quad [\mathrm{T}], \qquad (3)$$

gde je N_2 broj navojaka indukcionog namotaja i A je površina poprečnog preseka jezgra.

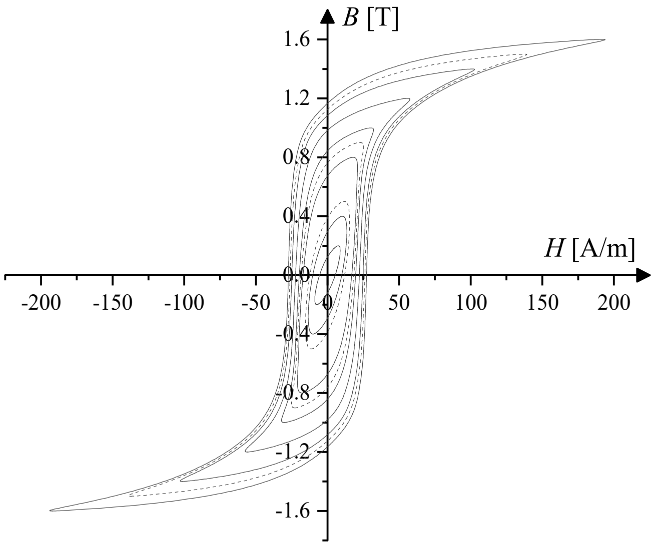
Podaci o broju pobudnih i indukcionih navojaka, dužini srednje linije i površini poprečnog preseka jezgra su dati u Tabeli I.

TABELA I
PODACI O TORUSNOM UZORKU

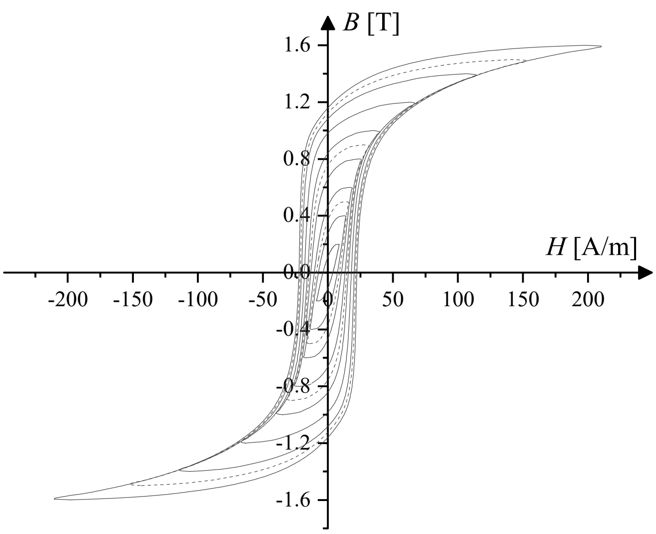
Veličina	Vrednost
N_1	175
N_2	60
l [m]	0,306
A [mm^2]	102,80

IV. REZULTATI MERENJA

Rezultati merenja su prikazani u obliku histerezisnih petlji. Na Sl. 2 su prikazane histerezisne petlje dobijene merenjima pri kontrolisanoj magnetskoj indukciji sinusnog i trougaonog oblika, respektivno. Merenja su izvršena pri frekvenciji od 50 Hz za opseg magnetskih indukcija amplituda od 0,2 T do 1,6 T sa korakom od 0,2 T. Takođe, isprekidanom linijom su prikazane i histerezisne petlje za verifikaciju simulacije za amplitude magnetske indukcije od 0,5 T, 0,9 T i 1,5 T.



a)



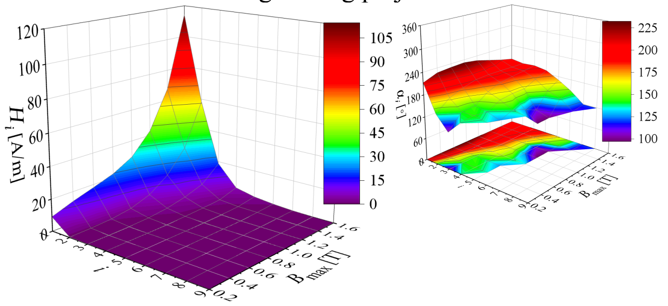
b)

Sl. 2. Izmerene histerezisne petlje pri a) sinusnom i b) trougaonom obliku magnetske indukcije.

V. REZULTATI PRORAČUNA I ANALIZA

Na Sl. 3 i 4 su prikazani rezultati proračuna harmonijskih komponenti magnetskog polja dobijenih merenjima.

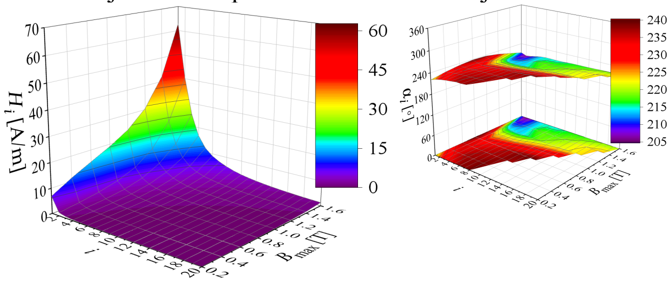
Na Sl. 3 je prikazana promena amplituda i faza harmonika magnetskog polja za magnetsku indukciju oblika sinusoide. Od prikazanih N=9 amplituda, koliko je od značaja za veće amplitude magnetske indukcije, samo 3 do 5 amplituda ima vrednost veću od 1 % od osnovnog harmonika. Shodno tome, radi preglednijeg prikaza, sa Sl. 3*b* su ukonjene faze viših harmonika, koje značajno osciluju, a koje nemaju veći uticaj na konačnu vrednost magnetskog polja.



a) *b*)

Sl. 3. Promena *a*) amplituda i *b*) faza harmonika magnetskog polja pri sinusnom obliku magnetske indukcije.

Na Sl. 4 je prikazana promena amplituda i faza harmonika magnetskog polja za magnetsku indukciju trougaonog oblika. Prikazano je N=20 amplituda i redukovani broj faza harmonika.



a) *b*)

Sl. 4. Promena *a*) amplituda i *b*) faza harmonika magnetskog polja pri trougaonom obliku magnetske indukcije.

Proračun harmonijskih komponenti magnetskog polja metodom interpolacije izvršen je za amplitude magnetske indukcije od 0,5 T, 0,9 T i 1,5 T. Usvojena vrednost najvišeg reda neparnih harmonika je N=15, za magnetsku indukciju sinusnog, i N=75, za magnetsku indukciju trougaonog oblika. Vrednosti za N usvojene su prema kriterijumu najmanjih oscilacija u vremenskom obliku magnetskog polja izračunatog primenom (1). Povećanje vrednosti N iznad usvojenih ne doprinosi značajnijem poboljšanju rezultata, a zadovoljavajući rezultati (sa minimalnim oscilacijama) se postižu i sa dva puta manjim brojem harmonika. Proračun harmonika je izvršen u programskom paketu *Mathematica* primenom funkcije *Fourier*, a interpolacija primenom funkcije *Interpolation*, čijom primenom se ceo postupak sprovodi bez dodatnog programiranja složenih matematičkih proračuna. Vremenski oblik magnetske indukcije je simuliran u istom programu kao idealna sinusna ili trougaona funkcija. Svaki signal je predstavljen nizom vrednosti sa ukupno 1000 članova niza. Deo korišćenog programskog koda koji pokazuje kako su određeni harmonici polja i kako je izvršena njihova interpolacija je dat u Dodatku.

U Tabelama II i III date su vrednosti prvih N=5 amplituda harmonika magnetskog polja dobijenih interpolacijom $H_{i,\mathrm{i}}$ i njihova odstupanja $\Delta H_{i\text{-}0,5}$, $\Delta H_{i\text{-}0,9}$ i $\Delta H_{i\text{-}1,5}$ od izračunatih amplituda harmonika izmerenih magnetskih polja H_i, gde je $\Delta H_i = H_{i,\mathrm{i}} - H_i$. Amplitude i odstupanja su dati za oba razmatrana oblika magnetske indukcije. Na osnovu rezultata prikazanih u Tabelama II i III je zaključeno da opisani postupak interpolacije harmonika ima zadovoljavajuću tačnost. Shodno tome, moguće je izvršiti simulaciju magnetskog polja korišćenjem izraza (1).

TABELA II

AMPLITUDE HARMONIKA MAGNETSKOG POLJA DOBIJENE INTERPOLACIJOM I NJIHOVA ODSTUPANJA OD AMPLITUDA DOBIJENIH IZRAČUNAVANJEM PRI MAGNETSKOJ INDUKCIJI SINUSNOG OBLIKA

i	$H_{i,\mathrm{i}\text{-}0,5}$ [A/m]	$H_{i,\mathrm{i}\text{-}0,9}$ [A/m]	$H_{i,\mathrm{i}\text{-}1,5}$ [A/m]	$\Delta H_{i\text{-}0,5}$ [A/m]	$\Delta H_{i\text{-}0,9}$ [A/m]	$\Delta H_{i\text{-}1,5}$ [A/m]
1	16,08	25,42	86,78	0,00	−0,07	0,09
2	0,86	4,93	41,51	0,00	−0,05	−0,03
3	0,16	1,29	12,23	0,02	−0,02	−0,07
4	0,03	0,27	2,84	0,01	0,02	−0,08
5	0,02	0,10	1,38	0,01	0,02	−0,03

TABELA III

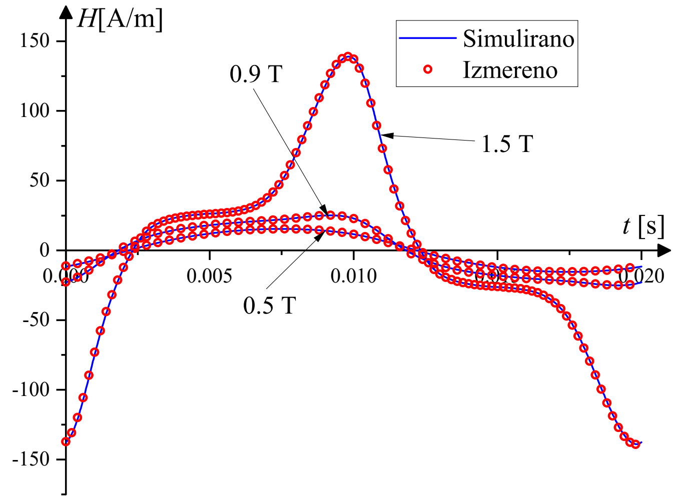
AMPLITUDE HARMONIKA MAGNETSKOG POLJA DOBIJENE INTERPOLACIJOM I NJIHOVA ODSTUPANJA OD AMPLITUDA DOBIJENIH IZRAČUNAVANJEM PRI MAGNETSKOJ INDUKCIJI TROUGAONOG OBLIKA

i	$H_{i,\mathrm{i}\text{-}0,5}$ [A/m]	$H_{i,\mathrm{i}\text{-}0,9}$ [A/m]	$H_{i,\mathrm{i}\text{-}1,5}$ [A/m]	$\Delta H_{i\text{-}0,5}$ [A/m]	$\Delta H_{i\text{-}0,9}$ [A/m]	$\Delta H_{i\text{-}1,5}$ [A/m]
1	13,36	21,18	50,94	0,13	−0,04	0,01
2	2,95	6,83	30,65	0,02	−0,05	−0,03
3	1,45	3,71	18,95	0,01	−0,04	−0,04
4	0,93	2,37	12,17	0,01	−0,02	−0,04
5	0,67	1,67	8,40	0,01	−0,01	−0,07

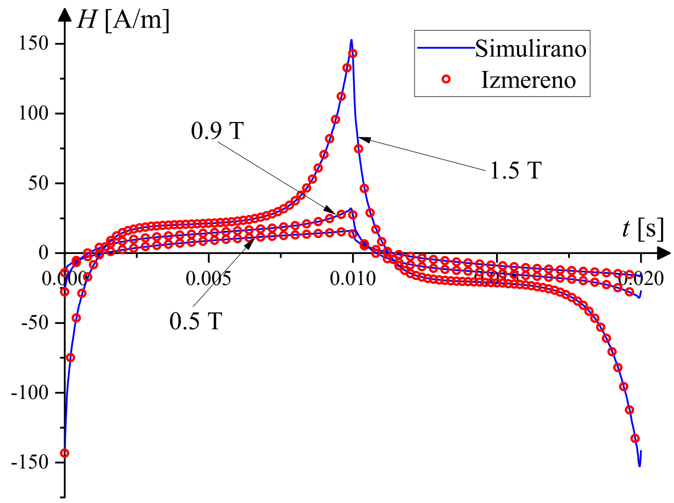
Na Sl. 5 i 6 prikazana su poređenja vremenskih oblika simuliranih i izmerenih magnetskih polja za razmatrane amplitude magnetske indukcije sinusnog i trougaonog oblika, respektivno. Mogu se uočiti dobra slaganja simuliranih i izmerenih magnetskih polja za oba razmatrana slučaja. Blaga odstupanja se uočavaju u trenucima kada magnetsko polje menja znak.

Kako je za merenje i proračun magnetskog polja korišćeno ukupno 1000 podataka, a slaganje izmerenih i simuliranih signala je veoma dobro, radi bolje preglednosti rezultata je simulirano polje prikazano punom linijom, a izmereno polje je prikazano tačkama, uz redukovan broj tačaka (jer 1000 tačaka takođe nije pregledno). Smanjenje broja tačaka je izvršeno bez brisanja podataka pomoću ugrađene opcije Skip Points u programu OriginLab. Na ovaj način je postignut željeni vizuelni efekat, bez uticaja na tačnost rezultata. Broj tačaka je odabran u skladu sa amplitudom polja.

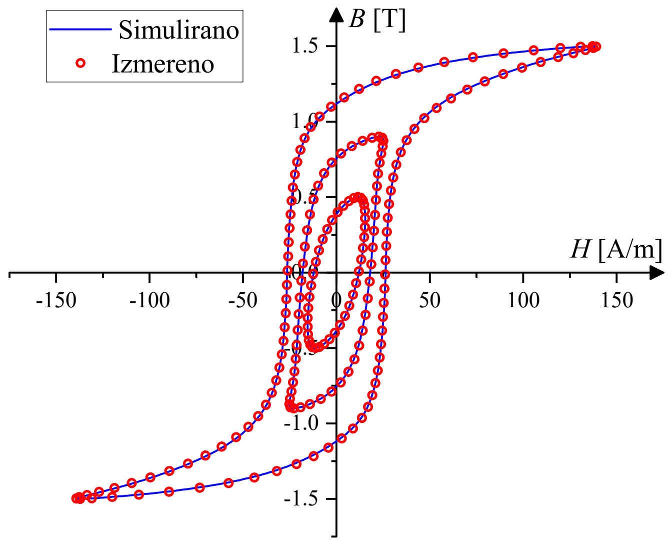
Na Sl. 7 i 8 prikazano je poređenje simuliranih i izmerenih histerezisnih petlji (zbog preglednosti rezultata, prikazan je manji broj tačaka za izmerene petlje). Uočava se dobro slaganje svih prikazanih petlji. Minimalna odstupanja simuliranih petlji mogu se pripisati kako odstupanjima kod magnetskih polja tako i odstupanjima izmerenih i simuliranih (idealnih) oblika magnetske indukcije.



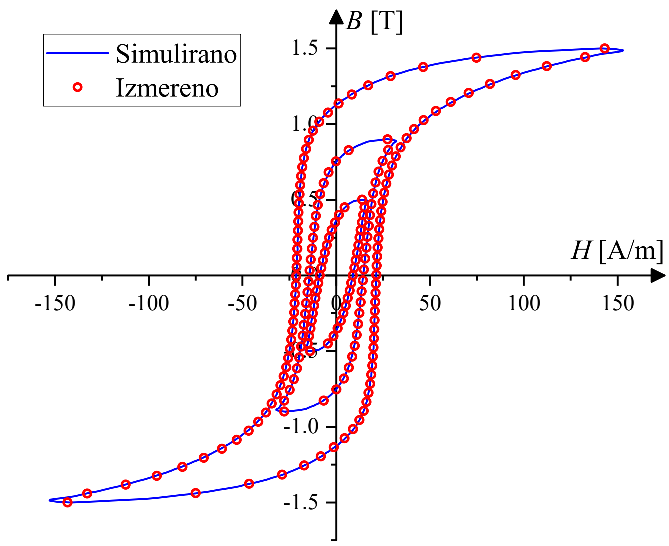
Sl. 5. Poređenje simuliranih i izmerenih trenutnih vrednosti magnetskog polja pri sinusnom obliku magnetske indukcije.



Sl. 6. Poređenje simuliranih i izmerenih trenutnih vrednosti magnetskog polja pri trougaonom obliku magnetske indukcije.



Sl. 7. Poređenje simuliranih i izmerenih histerezisnih petlji pri sinusnom obliku magnetske indukcije.



Sl. 8. Poređenje simuliranih i izmerenih histerezisnih petlji pri trougaonom obliku magnetske indukcije.

Dodatno, radi provere tačnosti kompletnog simulacionog postupka izvršeno je izračunavanje površina histerezisnih petlji i njihova analiza. U Tabeli IV je prikazano poređenje relativnih odstupanja površina histerezisnih petlji dobijenih simulacijom S_{sim} i merenjem S_{izm} za sinusni, δS_s, i trougaoni, δS_t, oblik indukcije, gde je δS [%]=100($S_{\mathrm{sim}}-S_{\mathrm{izm}}$)/$S_{\mathrm{izm}}$.

TABELA IV
POREĐENJE RELATIVNIH ODSTUPANJA POVRŠINA HISTEREZISNIH PETLJI

$B_{\max}$ [T]	δS_s [%]	δS_t [%]
0,5	−0,24	0,06
0,9	−1,50	−0,71
1,5	−1,21	0,69

Relativna odstupanja površina petlji su mala – reda veličine procenta i za sinusni i za trougaoni oblik magnetske indukcije.

Može se zaključiti da je tačnost predloženog postupka simulacije histerezisnih petlji zadovoljavajuća u svim razmatranim aspektima, pa se može očekivati i zadovoljavajuća tačnost u primeni istog u složenijim proračunima sa magnetskim kolima. Primenom prikazane harmonijske analize i postupka simulacije pri različitim frekvencijama pobudnog polja može se dobiti kompletnija slika o karakteristikama magnetskog materijala ili jezgra. Dodatno, ovaj postupak omogućava rešavanje i analizu magnetskih problema u vremenskom domenu.

VI. ZAKLJUČAK

U ovom radu je prikazan postupak simulacije histerezisnih petlji za određene amplitude i oblike magnetske indukcije. Postupak se zasniva na interpolaciji amplituda i faza harmonika, a zatim i simulaciji magnetskog polja i magnetske indukcije od interesa.

U radu je ukratko opisana metoda merenja magnetskih karakteristika i predstavljene su izmerene histerezisne petlje za sinusni i trougaoni oblik magnetske indukcije sa amplitudama od 0,2 T do 1,6 T sa korakom od 0,2 T. Rezultati harmonijske analize odgovarajućih magnetskih polja su takođe prikazani i analizirani. Na osnovu tako dobijenih harmonika je izvršena simulacija novih magnetskih polja i histerezisnih petlji. Rezultati simulacije su analizirani kroz poređenje interpoliranih i izračunatih amplituda harmonika, trenutnih vrednosti magnetskog polja, oblika histerezisnih petlji i njihovih površina. Utvrđeno je vrlo dobro slaganje svih simuliranih i izmerenih rezultata.

Na osnovu prikazane analize rezultata je zaključeno da se opisani postupak može primeniti za kvalitetniju karakterizaciju feromagnetskih materijala i za rešavanje problema sa magnetskim kolima u vremenskom domenu.

DODATAK

Deo programskog koda za učitavanje merenih podataka, simulaciju magnetske indukcije, proračun harmonika magnetskog polja i njihovu interpolaciju je prikazan na Sl. 9. Prikazan je najvažniji deo koda u kojem se vidi način primene funkcija *Fourier* i *Interpolation*. Ostatak koda za proračun veličina od interesa i njihov numerički ili grafički prikaz je izostavljen jer prevazilazi obim rada, pri čemu su ti rezultati već prikazani u samom radu.

```
Unos podataka - Sinus

Podaci = Import["F:\\Desktop\\RADOVI\\RADOVI ETRAN 2021\\Divac\\50Hz-Sin-Loops.txt", "Table"];
PodaciProvere = Import["F:\\Desktop\\RADOVI\\RADOVI ETRAN 2021\\Divac\\50Hz-1_5T-Sin.txt", "Table"];

Bzeljeno = 1.5;
Nh = 5; (*Broj neparnih harmonika*)

Simulacija B

Bvr = Table[Bzeljeno * Sin[2 * 50 * Pi * t + 3 * Pi / 2], {t, 0, (0.02 - 0.02 / 1000), (0.02 / 1000)}];

Interpolacija - Pojedinacni Harmonici

PodaciPolja = Table[Podaci[[All, i]], {i, 2, Length[Podaci[[1, All]]], 2}];
PodaciIndukcije = Table[Podaci[[All, i]], {i, 3, Length[Podaci[[1, All]]], 2}];
PodaciPolja = Transpose[PodaciPolja];
PodaciIndukcije = Transpose[PodaciIndukcije];

AmplitudeHarmonika = Table[0, {i, Nh}, {j, Length[PodaciPolja[[1, All]]]}];
FazeHarmonika = Table[0, {i, Nh}, {j, Length[PodaciPolja[[1, All]]]}];

For[j = 1, j ≤ Length[PodaciPolja[[1, All]]], j++, HVr = Table[{Podaci[[i, 1]], PodaciPolja[[i, j]]}, {i, 1000}];
 ft = Fourier[HVr[[All, 2]], FourierParameters → {-1, -1}];
 pom1 = 2 * Abs[ft];
 pom2 = Arg[ft] * 180 / Pi;
 FazPolja = Table[pom2[[i]], {i, 2, Nh * 2, 2}];
 APolja = Table[pom1[[i]], {i, 2, Nh * 2, 2}];
 For[i = 1, i < Nh, i++, If[FazPolja[[i]] < 0, FazPolja[[i]] = FazPolja[[i]] + 360, FazPolja[[i]] = FazPolja[[i]]]];
 For[i = 1, i ≤ Nh, i++, AmplitudeHarmonika[[i, j]] = APolja[[i]];
  FazeHarmonika[[i, j]] = FazPolja[[i]]]]

AmplitudeHarmonika = Transpose[AmplitudeHarmonika]
FazeHarmonika = Transpose[FazeHarmonika]

Bmax = Table[Max[PodaciIndukcije[[All, i]]], {i, Length[PodaciIndukcije[[1, All]]]}];

Ampl = Table[0, {i, Nh}];
Faz = Table[0, {i, Nh}];

For[k = 1, k <= Nh, k++, PodaciInterpolacijeAmplituda = Table[{Bmax[[i]], AmplitudeHarmonika[[i, k]]}, {i, 1, Length[Bmax]}];
 PodaciInterpolacijeFaza = Table[{Bmax[[i]], FazeHarmonika[[i, k]]}, {i, 1, Length[Bmax]}];
 FunkcijaA = Interpolation[PodaciInterpolacijeAmplituda, InterpolationOrder → 6];
 FunkcijaF = Interpolation[PodaciInterpolacijeFaza, InterpolationOrder → 1];
 Ampl[[k]] = FunkcijaA[Bzeljeno];
 Faz[[k]] = FunkcijaF[Bzeljeno];]
(*--------------------------------------------------------*)
```

Sl. 9. Deo programskog koda za simulaciju magnetskog polja i indukcije.

ZAHVALNICA

Istraživanja prezentovana u ovom radu su delimično finansirana sredstvima Ministarstva prosvete, nauke i tehnološkog razvoja RS, Ugovor br. 451-03-9/2021-14/200132, čiji je realizator Fakultet tehničkih nauka u Čačku - Univerziteta u Kragujevcu.

LITERATURA

[1] C. Grech, M. Buzio, M. Pentella, M. Sammut, "Dynamic Ferromagnetic Hysteresis Modelling Using a Preisach-Recurrent Neural Network Model", *Materials.*, vol. *13*, no. *11*, Spec. issue: *Model. and Char. of Magn. Mat.*, June 2020.

[2] S. E. Zirka, Y. I. Moroz, R. G. Harrison, K. Chwastek, "On physical aspects of the Jiles-Atherton hysteresis models", *J. of Appl. Phys.*, vol. *112*, no. *8*, Aug. 2012.
[3] Y. Li, R. Chen, Z. Cheng, C. Zhang, L. Liu, "Dynamic Hysteresis Lops Modeling of Electrical Steel with Harmonic Components", *IEEE Trans. on Ind. Appl.*, vol. *56*, no. *5*, pp. 4804-4811, Sept.-Oct. 2020.
[4] A. Milovanović, B. Koprivica, "Mathematical Model of Major Hysteresis Loop and Transient Magnetizations", *Electromagnetics*, vol. *35*, no. *3*, pp. 155-166, Match 2015.
[5] Y. Hane, K. Nakamura, "Dynamic Hysteresis Modeling for Magnetic Circuit Analysis by Incorporating Play Model and Cauer's Equivalent Circuit Theory", *IEEE Trans. on Magn.*, vol. *56*, no. *8*, Aug. 2020.
[6] Q. Antonio, R. Fulginei, H. P. Rimal, A. M. Ghanim, "On the Use of Feedforward Neural Networks to Simulate Magnetic Hysteresis in Electrical Steelsy", *20th Mediterranean Electrotechnical Conference (MELECON)*, Palermo, Italy, 16th-18th June 2020.
[7] S. Q. Antonio, F. R. Fulginei, K. Nakamura, A. Laudani, A. Faba, E. Cardelli, "An Effective Neural Network Approach to Reproduce Magnetic Hysteresis in Electrical Steel under Arbitrary Excitation Waveforms", *J. of Magn. and Magn. Mat.*,vol. *528*, June 2021.
[8] G. Bramerdorfer, D. Andessner, "Accurate and Easy to Obtain Iron Loss Model for Electric Machine Designs", *IEEE Trans. on Indu. Electr.*, vol. *64*, no. *3*, pp. 2530-2537, March 2017.
[9] B. Koprivica, I. Dumitru, A. Milovanović and O. Caltun, "Harmonic Analysis and Modelling of Magnetisation Process in Soft Ferromagnetic Material", *FACTA Universitatis, Electronics and Energetics*, vol. *30*, no. *1*, pp. 121 – 136, March 2017.
[10] B. Koprivica, A. Milovanović and M. Đekić, "Effects of Wound Toroidal Core Dimensional and Geometrical Parameters on Measured Magnetic Properties of Electrical Steel", *Serb. J. of Electr. Eng.*, vol. *10*, no. *3*, pp. 459-471, Oct. 2013.
[11] W. Cheng and Y. Saito, "Modeling and Analysis pf Hysteresis by Harmonic Balance Method", *J. of Appl. Phys.*, vol. *117*, no. *17*, May 2015.
[12] S. Zurek, P. Marketos, T. Meydan, and A.J. Moses, "Use of Novel Adaptive Digital Feedback for Magnetic Measurements Under Controlled Magnetizing Conditions", *IEEE Trans. on Magn.*, vol. *41*, no. *11*, pp. 4242-4249, Nov. 2005.
[13] J. V. Surutka, "*Elektromagnetika*", *osmo izdanje*, Akademska misao, Beograd, 2006.

ABSTRACT

The aim of this paper is to present a procedure for simulation of hysteresis loops of a ferromagnetic sample which is performed by simulation of waveforms of the magnetic field strength (further magnetic field) and flux density. The data for magnetic field and flux density are obtained by measurement, under controlled shape of magnetic flux density, for two considered shapes - sinusoidal and triangular. The harmonics of the magnetic field for the known amplitudes of the magnetic flux density were determined according to these data. New values of these harmonics are calculated by interpolating ones previously calculated, for the amplitude of magnetic flux density of interest. Simulation of the new magnetic field waveform is performed using these harmonics. The new hysteresis loop is simulated using this magnetic field and new simulated magnetic flux density. Description of this simulation procedure, comparison of calculated and measured amplitudes of harmonics and magnetic field waveforms, as well as hysteresis loops, is given in this paper. Appropriate analysis and discussion of the results is also given.

Simulation of hysteresis loops by interpolation of harmonic components of magnetic field

Srđan Divac, Branko Koprivica

Analiza uticaja magnetske interakcije faza na karakteristike 8/6 SRM-a

Dragan S. Mihić, *Member, IEEE*, Mladen V. Terzić, *Member, IEEE,* Žarko V. Koprivica, *Student, IEEE,* Bogdan M. Brković, *Member, IEEE*

Apstrakt— **U cilju unapređenja performansi, rad prekidačkog reluktantnog motora (Switched Reluctance Motor-SRM) zahteva jednovremeno pobuđivanje više faza motora. Budući da magnetski polaritet pobuđenih faza motora može biti isti ili različit, u ovom radu je izvršena detaljna analiza uticaja magnetskog polariteta na fluksne obuhvate faza 8/6 SRM-a. Ova analiza je izvršena na osnovu statičkih karakteristika koje su dobijene korišćenjem modela baziranog na metodi konačnih elemenata (Finite Element Method-FEM) u softverskom paketu Ansys Electronics. Pored toga, imajući u vidu da kod 8/6 SRM-a korišćenjem standardnog asimetričnog polumostnog invertora (API) nije moguće ostvariti magnetsku simetriju faza, rezultujući efekti veoma izražene interakcije faza su analizirani na primeru odgovarajućih tranzijentnih talasnih oblika faznih struja i momenta.**

***Ključne reči*—Prekidački reluktantni motor, interakcija faza, metoda konačnih elemenata.**

I. Uvod

Prekidački reluktantni motori zbog svojih istaknutih osobina kao što su jednostavna i jeftina proizvodnja, mehanička i termička robusnost, velika gustina snage, širok opseg brzina i pouzdanost, privlače značajnu pažnju i postaju značajni konkurenti mašinama naizmenične struje pre svega u električnim automobilima, vetrogeneratorima, avio industriji [1-4].

U cilju unapređenja efikasnosti, rad SRM-a zahteva jednovremeno pobuđivanje više faza motora [5]. Na taj način se doprinosi uvećanju momenta i smanjenu njegove valovitosti. To je naročito izraženo za konfiguracije SRM-a sa brojem faza koji je veći od tri. Broj istovremeno pobuđenih faza zavisi od radnog režima i u opštem slučaju taj broj može biti veći od dva. Interakcija faza može biti rezultat istovremene magnetizacije faze i demagnetizacije njoj susedne faze ili istovremene magnetizacije susednih faza motora kada zbog velike širine ugla vođenja dolazi i do značajnije interakcije ortogonalnih faza SRM-a. Kao posledica istovremenog pobuđivanja faza mogu se javiti različite putanje magnetskog fluksa u mašini. Karakter i specifičnosti tih putanja su vezane za magnetski polaritet pobuđenih faza koji zavisi od smera struje u njima. Budući da magnetski polaritet pobuđenih faza može biti isti ili različit, tokom rada motora, raspodela magnetskog fluksa u mašini je direktno određena njihovim trenutnim stanjem. Dakle, međusobni fluks kao i rezultujuće magnetsko zasićenje koje se javlja usled interakcije faza je direktno vezano za magnetski polaritet pobuđenih faza motora čime se direktno utiče na performanse SRM-a.

Da bi se utvrdili efekti interakcije faza na fluksne obuhvate a samim tim i na performanse mašine neophodno je korišćenje tačnog modela koji uzima u obzir efekte međusobne sprege i zasićenja tokom istovremenog vođenja faza. Generalno, poteškoće koje se javljaju tokom modelovanja i analize SRM-a su posledica izrazito nelinearne veze fluksa, struje i položaja rotora budući da je normalan rad SRM-a praćen dubokim zasićenjem. Zbog toga, zahtev za modelovanjem efekata međusobne interakcije faza predstavlja dodatnu poteškoću.

U [6] je analiziran uticaj polariteta sukcesivnih faza motora sa konstantnim strujama na raspodelu fluksa u različitim delovima mašine i generisanje elektromagnetskog momenta. Navedene su neke kvalitativne osobine dugih i kratkih putanja fluksa koje mogu biti osnova za unapređenje tehnika koje se bave smanjenjem valovitosti momenta i gubitaka u gvožđu. Međutim, nisu prikazani rezultati tranzijentnih simulacija koji prikazuje uticaj efekata interakcija faza na fluksne obuhvate, fazne struje i momenat.

Mnogi radovi [7-12] naglašavaju važnost modelovanja međusobne sprege faza i njenog uticaja na tranzijentne karakteristike SRM-a. Međutim, u odgovarajućim razvijenim matematičkim modelima uticaj struje koja postoji u faznom namotaju na odgovarajući fluksni obuhvat tokom interakcije sa drugim pobuđenim fazama se zanemaruje pa zbog toga ovi efekti i nisu analizirani.

Dinamički modeli [13-15] uzimaju u obzir međusobnu spregu faza bazirajući se na karakteristikama magnetostatičkog FEM-a ili eksperimenta. Pri tome, ulazne statičke karakteristike razmatraju uticaj samo dve pobuđene faze što je ograničavajuće u analizi i modelovanju dinamičkih radnih režima kada su istovremeno pobuđene tri ili četiri faze motora, što se uobičajeno ima u slučaju četvorofaznih SRM-a. Nelinearni analitički model [16] uzima u obzir efekte interakcije faza, ali njegove tranzijentne simulacije u opštem slučaju ne pružaju mogućnost da se u potpunosti, na eksplicitan način, analizira uticaj rezultujućeg magnetskog zasićenja na karakteristike motora tokom istovremenog

Dragan S. Mihić – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: dragan84m@ etf.bg.ac.rs).

Mladen V. Terzić – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: terzic@ etf.bg.ac.rs).

Žarko V. Koprivica – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: kopri@ etf.bg.ac.rs).

Bogdan M. Brković – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: brkovic@ etf.bg.ac.rs).

vođenja više faza motora.

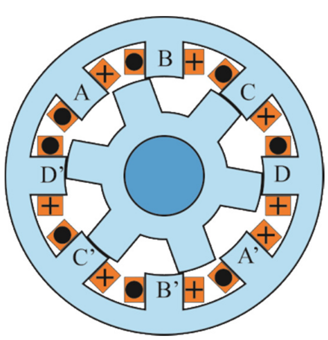
U ovom radu je izvršena detaljna analiza uticaja magnetskog polariteta na fluksne obuhvate faza 8/6 SRM-a. Ova analiza je izvršena na osnovu statičkih karakteristika koje su dobijene korišćenjem modela baziranog na FEM-u u softverskom paketu Ansys Electronics. Pored toga, imajući u vidu da kod 8/6 SRM-a korišćenjem standardnog asimetričnog polumostnog invertora nije moguće ostvariti magnetsku simetriju faza, rezultujući efekti veoma izražene interakcije faza su prikazani i analizirani na primeru odgovarajućih tranzijentnih talasnih oblika faznih struja i momenta.

U II poglavlju su navedeni glavni koraci u modelovanju 8/6 SRM-a metodom konačnih elemenata u programu Ansys Electronics u cilju dobijanja statičkih i dinamičkih karakteristika. Analiza uticaja magnetskog polariteta na statičke karakteristike fluksnih obuhvata 8/6 SRM-a je data u poglavlju III. Izraženi rezultujući efekti interakcije faza na dinamičke talasne oblike faznih struja i elektromagnetskog momenta su prikazni i analizirani u poglavlju IV. Zaključak je dat u V poglavlju.

II. Modelovanje 8/6 SRM-a metodom konačnih elemenata u programskom paketu Ansys Electronics

Na Sl. 1 je prikazan poprečni presek 8/6 SRM-a za koji se određuju statičke i dinamičke karakteristike u cilju utvrđivanja uticaja magnetskog polariteta faza na njegove karakteristike. Glavni parametri i dimenzije 8/6 SRM-a su dati u Tabelama 1 i 2, respektivno. Magnetostatički i tranzijentni FEM model, uključujući 8/6 SRM geometriju, materijale, električno i magnetsko kolo, granične uslove, mrežu konačnih elemenata u programskom paketu Ansys Electronics su definisani saglasno [16]. Pored toga, da bi se dobile statičke karakteristike fluksnih obuhvata, definisan je odgovarajući opseg konstantnih struja posmatranih faza motora na ugaonom intervalu od neusaglašenog do usaglašenog položaja rotora. Sa druge strane, za dobijanje rezultata tranzijentne simulacije definisano je eksterno kolo koje predstavlja API-a. Gore navedeni model je napravljen uzimajući u obzir da znak “×” predstavlja provodnike u kojima postoji struja u smeru od posmatrača, dok oznaka “● ”predstavlja provodnike u kojima struja postoji u smeru ka posmatraču.

U skladu sa ovim, primenom unipolarnog napajanja rezultujuća magnetska sekvenca u pobuđivanju faza A-B-C-D-A-B-C-D je S-S-S-S-N-N-N-N. Kod ove magnetske sekvence tri od četiri kombinacije susednih faza motora su istog, a jedna različitog magnetskog polariteta.



Sl.1. Poprečni presek 8/6 SRM-a

TABELA 1
GLAVNI PARAMETRI 8/6 SRM-A

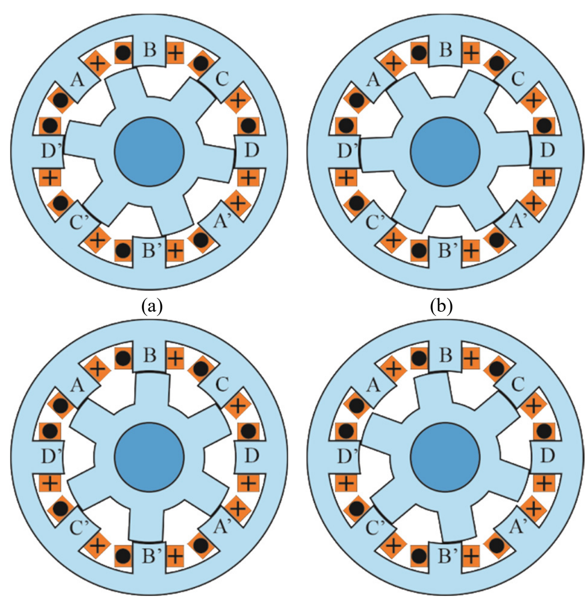
Parametri	Vrednost
Broj faza	4
Broj statorskih/rotorskih polova	8/6
Nominalni fazni napon [V]	220
Nominalna fazna struja [A]	3.2
Fazna otpornost [Ω]	2.1
Induktivnost u neusaglašenom položaju [mH]	14
Broj navojaka po fazi	284

TABELA 2
GLAVNE DIMENZIJE 8/6 SRM-A

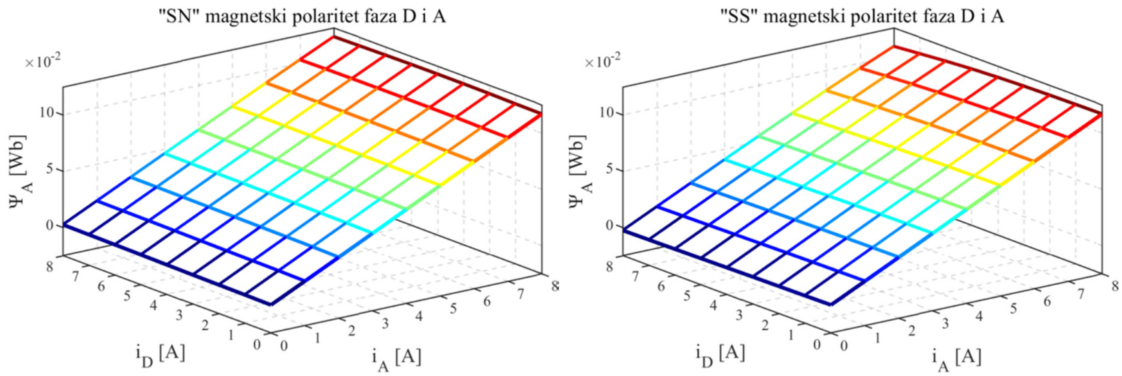
Dimenzije	Vrednost
Spoljašnji poluprečnik statora [mm]	60
Unutrašnji poluprečnik statora [mm]	37.5
Debljina jarma statora [mm]	9
Spoljašnji poluprečnik rotora [mm]	37
Unutrašnji poluprečnik rotora [mm]	24
Debljina jarma rotora [mm]	9
Širina vazdušnog zazora [mm]	0.5
Poluprečnik vratila [mm]	15
Dužina magnetskog kola [mm]	65
Širina statorskog pola [°]	22
Širina rotorskog pola [°]	23

III. Analiza uticaja interakcije faza na statičke Ψ-I karakteristike

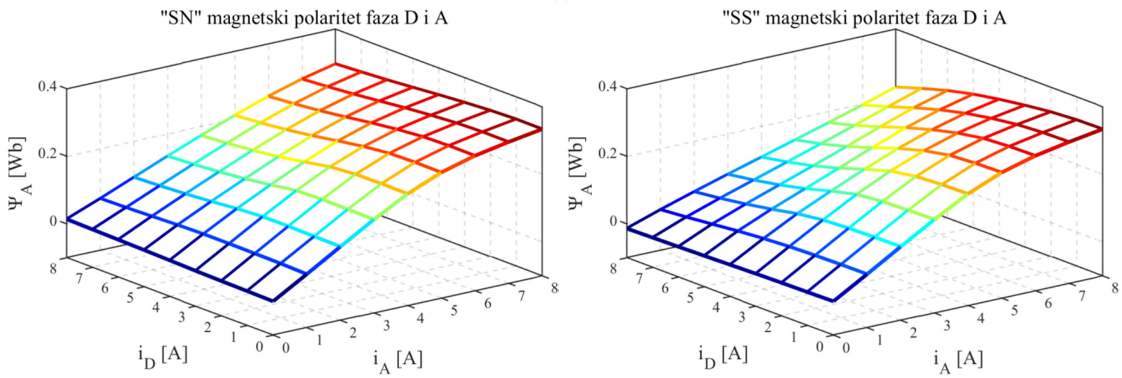
Da bi se utvrdio uticaj magnetskog polariteta faza na fluksni obuhvat kada je jaram statora i rotora, usled rezultujućih efekata interakcije faza u dubokom zasićenju, od interesa je analizirati odgovarajuće statičke Ψ-i karakteristike posmatrane faze tokom njene interakcije sa fazom koja joj prethodi, ali i fazom koja je sledi za različite ugaone položaje rotora. U suštini, potpuno je svejedno za koju fazu se analiziraju navedene karakteristike, pa će se u nastavku posmatrati Ψ-i karakteristike faze A kao i rezultujući efekti međusobne interakcije sa fazama D i B za različite ugaone položaje rotora. Od interesa je posmatrati takve ugaone položaje rotora koji će omogućiti da se utvrdi priroda uticaja



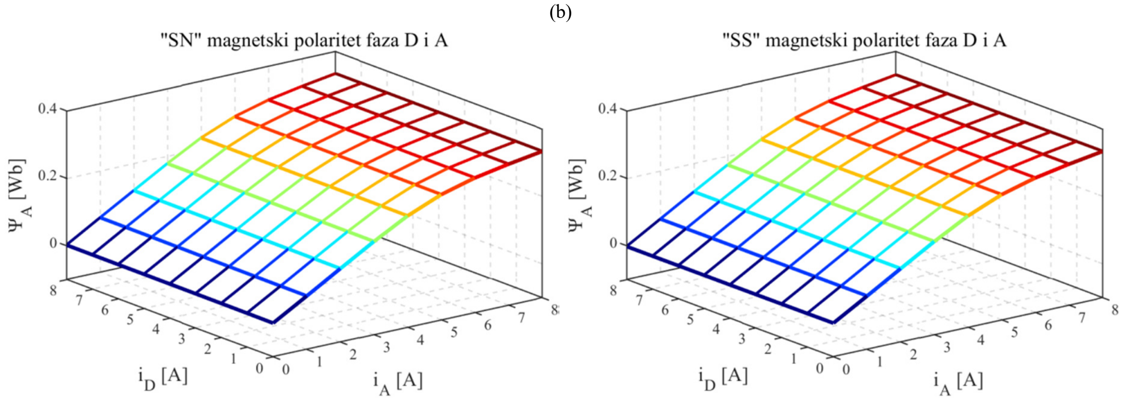
Sl. 2. Ugaoni položaji rotora (a) θ_m=5°, (b) θ_m=17.5°, (c) θ_m=42.5° i (d) θ_m=55° u odnosu na neusaglašeni položaj faze A.



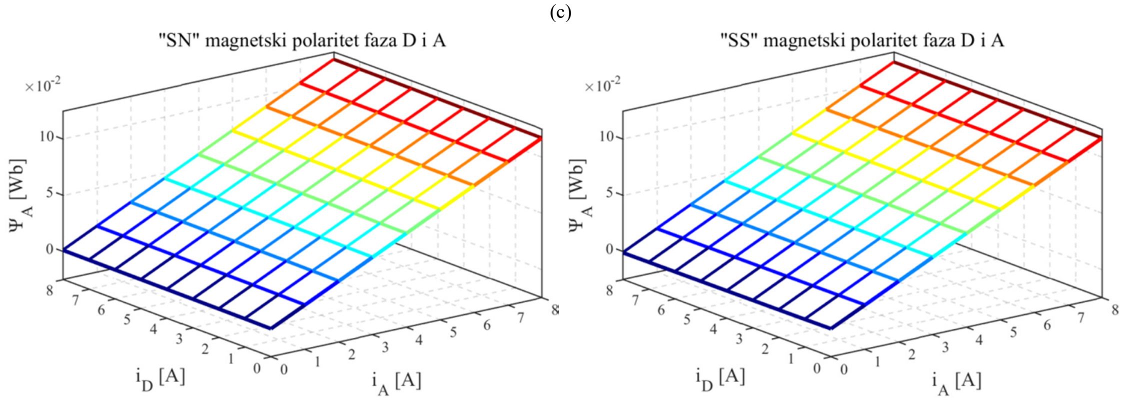
(a)



(b)



(c)



(d)

Sl. 3. Ψ-i karakteristike faze A za ugaoni položaj rotora (a) θ_m=5°, (b) θ_m=17.5°, (c) θ_m=42.5° и (d) θ_m=55° kada je magnetski polaritet susednih faza D i A različit (SN) i isti (SS).

međusobne interakcije na Ψ-i karakteristiku faze na celom opsegu polnog koraka rotora faze A. Zbog toga, u odnosu na referentni ugaoni položaj rotora koji odgovara neusaglašenom položaju faze A, izabrana su četiri ugaona položaja rotora θ_m

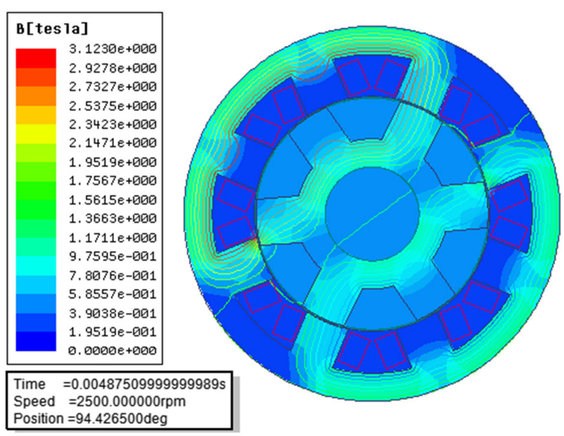
saglasno Sl. 2.

Statičke Ψ-i karakteristike faze A na Sl. 3, koje odgovaraju ugaonim položajima rotora θ_m=5°; θ_m=17.5°; θ_m=42.5° i θ_m=55°, respektivno, i opsegu struja i_A, i_D=(0,1,2,…,8) A, ukazuju na uticaj struje faze D na fluksni obuhvat faze A kada je njihov magnetski polaritet različit odnosno isti. Pri tome, od interesa je naglasiti da se promena magnetskog polariteta faze D u FEM modelu vrši promenom smera odgovarajuće struje.

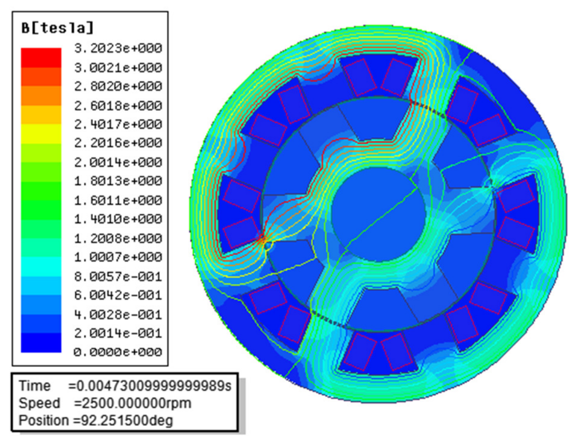
Sa povećanjem struje i_D pri konstantnoj vrednosti struje i_A, u slučaju različitog magnetskog polariteta faza A i D, fluksni obuhvat Ψ_A monotono raste do određene vrednosti usled pozitivnog međusobnog fluksa koji postoji između njih. Međutim, nakon određene vrednosti struje i_D on se smanjuje usled dubokog magnetskog zasićenja delova jarma statora i rotora gde postoji preklapanje flukseva ovih faza. Pri istim uslovima, u slučaju istog magnetskog polariteta, usled negativnog međusobnog fluksa, ima se monotono opadanje fluksnog obuhvata Ψ_A sa većom strminom opadanja kada jaram statora i rotora uđe u zasićenje. Dakle, statičke Ψ-i karakteristike jasno ukazuju na uticaj polariteta međusobnog fluksa na ukupni fluksni obuhvat faze.

Efekat opadanja fluksnog obuhvata kada nastupi duboko zasićenje je izraženiji u slučaju istog magnetskog polariteta budući da su tada putanje magnetskog fluksa u opštem slučaju duže. To se može jasno uočiti posmatrajući raspodelu magnetskog polja na Sl. 4 i 5. Na Sl. 4 su date raspodele magnetskog polja u trenucima kada su fluksni obuhvat faze A i D isti i različiti po amplitudi pri istom magnetskom polaritetu. Sa druge strane, Sl. 5 odgovara raspodeli magnetskog polja u istim trenucima, ali u slučaju različitog magnetskog polariteta. Kao posledica gore navedenog, veći pad magnetskog napona, odnosno manji fluksni obuhvat se ima u slučaju istog magnetskog polariteta nego u slučaju kada je on različit.

Dakle, uticaj faze koja prethodi posmatranoj se manifestuje na takav način da se ima veći fluksni obuhvat posmatrane faze kada je njihov magnetski polaritet različit. Pri tome, treba naglasiti da se benefit različitog magnetskog polariteta u odnosu na isti osetnije javlja prilikom veće preklopljenosti statorskih i rotorskih polova posmatrane faze A. Sa druge strane, kada ne postoji njihovo preklapanje ili se ono smanjuje, tada su efekti međusobne interakcije faza manje izraženi. Odnos fluksnih obuhvata, čak i pri dubokom zasićenju, je približno isti u oba slučaja. Međutim, uvek je veći u slučaju različitog magnetskog polariteta. Ove pojave se mogu jasno uočiti na Sl. 6, gde su za posmatrane ugaone položaje rotora izdvojene Ψ-i karakteristike faze A pri vrednosti struje faze D od 2 A i 8 A.

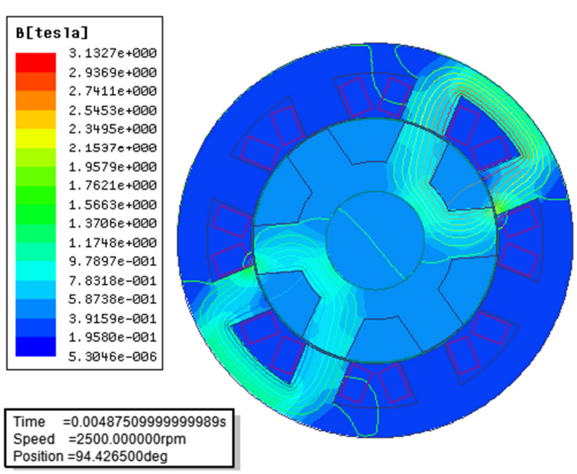


(a)

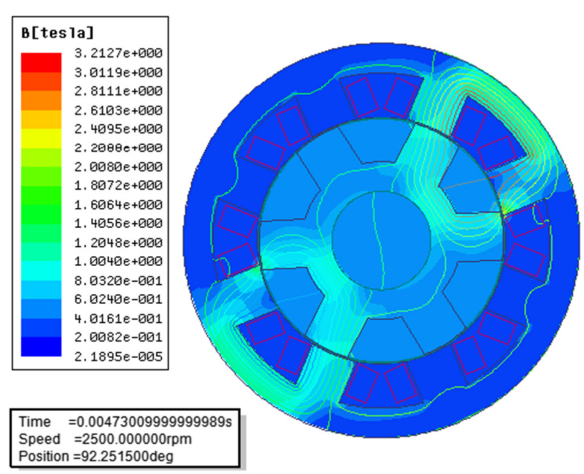


(b)

Sl. 4. Raspodela magnetskog polja u trenutku kada je (a) Ψ_D=Ψ_A i (b) Ψ_D>Ψ_A pri čemu su istovremeno pobuđene faze D i A istog magnetskog polariteta.

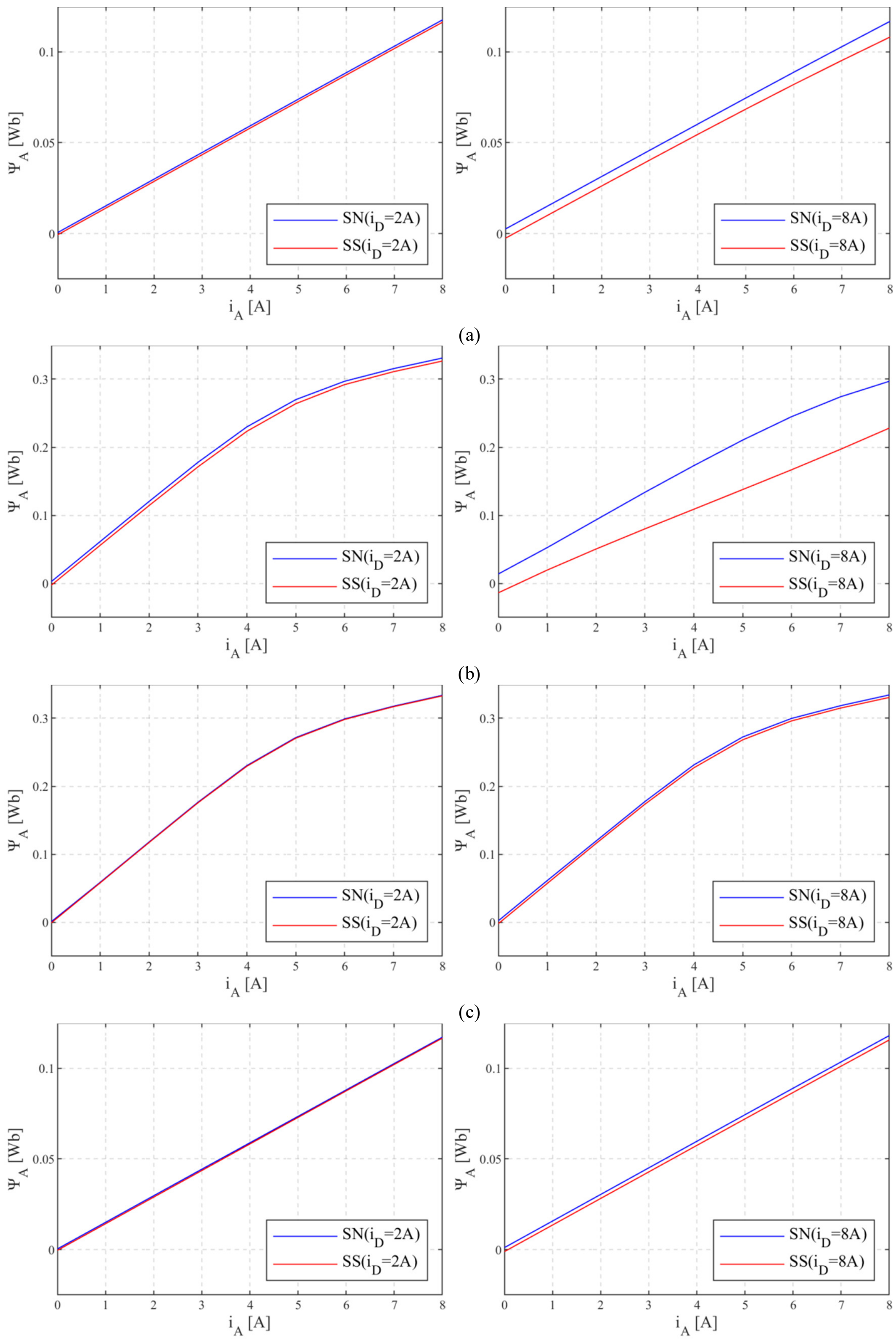


(a)



(b)

Sl. 5. Raspodela magnetskog polja u trenutku kada je (a) Ψ_D=Ψ_A i (b) Ψ_D>Ψ_A pri čemu su istovremeno pobuđene faze D i A različitog magnetskog polariteta.



(d)

Sl. 6. Uticaj struje faze D na fluksni obuhvat faze A za ugaoni položaj rotora (a) θ_m=5°, (b) θ_m=17.5°, (c) θ_m=42.5° i (d) θ_m=55° kada je magnetski polaritet susednih faza D i A različit (SN) i isti (SS).

Slično gore sprovedenoj analizi, potrebno je utvrditi i efekte interakcije koji se javljaju kao posledica istovremenog vođenja faza A i B. Do ovih zaključaka se može doći na osnovu rezultata koji prikazuju uticaj struje faze D na fluksni obuhvat faze A. Naime, pažljivom analizom se može zaključiti da raspodeli magnetskog fluksa u mašini, kada su istovremeno pobuđenje faze D i A, za neki ugaoni položaj rotora θ_m, odgovara ista raspodela magnetskog fluksa za ugaoni položaj rotora θ_m'=60-θ_m kada su istovremeno pobuđene faze A i B. Pri tome je od interesa naglasiti da se za definisane kombinacije susednih faza motora posmatra odgovarajući magnetski polaritet. Zbog toga se u položaju rotora θ_m=5° i θ_m'=55° respektivno ima ista Ψ-i karakteristika faze A kada su istovremeno pobuđene faze D i A odnosno A i B, respektivno. Slično važi i za druga dva razmatrana položaja rotora.

IV. Analiza uticaja interkcije faza na tranzijentne karakteristike

Prethodno sprovedena analiza koja se odnosila na statičke Ψ-i karakteristike pruža mogućnost da se analizira i uticaj magnetskog polariteta faza na tranzijentne karakteristike. To će se pokazati na primeru radne tačke koju karakteriše veoma izražena interakcija faza. Ova radna tačka je definisana kontrolnim parametrima: naponom napajanja (V_{DC}), uglom uključenja (θ_{ON}), uglom isključenja (θ_{OFF}) i brzinom obrtanja (n) koji su dati u Tabeli 3. Odgovarajući talasni oblici faznih struja i elektromagnetskog momenta su prikazani na Sl. 7 i 8, respektivno.

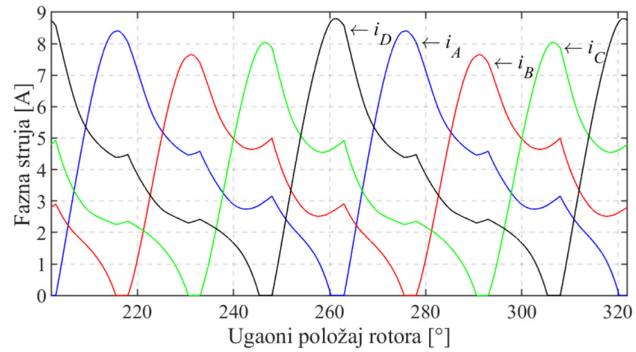
Saglasno Sl. 7, posmatranu radnu tačku karakteriše veliko preklapanje u vođenju faza budući da su tokom električnog ciklusa pobuđene tri ili sve četiri faze motora. Kao posledica izraženih efekata interakcije faza dolazi do nesimetrije faznih struja jer sekvencu pobuđivanja faza A-B-C-D-A-B-C-D karakteriše magnetska nesimetrija S-S-S-S-N-N-N-N. Posmatrajući magnetsku sekvencu, može se uočiti da je tokom magnetizacije kod sve tri faze koje prethode fazi D magnetski polaritet isti kao za fazu D. Zbog toga je i vršna vrednost struje i_D najveća. Budući da pobuđivanju faze A prethodi pobuđivanje faze D, saglasno tome i magnetskom polaritetu preostalih faza koje joj prethode, nakon faze D efekti interakcije faza su najizraženiji u fazi A. Sa druge strane, za faze B i C taj efekat je približno isti, ali je u izvesnoj meri izraženiji u fazi C jer se u dve od tri faze koje prethode fazi C ima isti magnetski polaritet kao kod faze C, dok se slučaju faze B to ima u jednoj od tri faze.

U trenutku početka demagetizacije faznog namotaja, efekti interakcije faza su najizraženiji u struji i_A jer je tada magnetski polaritet kod faze A isti kao kod preostalih faza koje je slede. Analogno, za razliku od magnetizacije, sada su ovi efekti izraženiji kod struje i_B u odnosu na struju i_C dok je najmanji uticaj na struju i_D. Pored razmatranih delova talasnih oblika faznih struja, od interesa je uočiti i pojavu koju karakteriše porast struje tokom demagnetizacije. Ova pojava je najviše izražena kod struje i_A a zatim kod i_B, i_D i i_C.

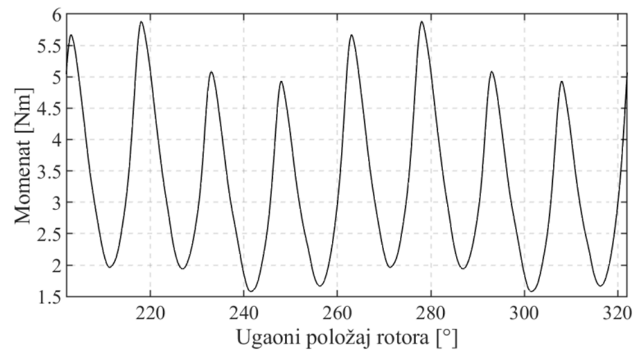
Izražena nesimetrija u talasnim oblicima faznih struja je u direktnoj vezi sa magnetskim polaritetom faza koje se isključuju ili uključuju i njihovim uticajem na preostale

TABELA 3
RADNI REŽIM SA IZRAŽENOM INTERAKCIJOM FAZA

V_{DC} [V]	θ_{ON} [°]	θ_{OFF} [°]	n [ob/min]
220	23	53	3500



Sl. 7. Fazne struje 8/6 SRM-a za radni režim definisan u Tabeli 3.



Sl. 8. Elektromagnetski momenat 8/6 SRM-a za radni režim definisan u Tabeli 3.

aktivne faze jer u trenutku komutacije faza dolazi do nagle promene rezultujućeg magnetskog zasićenja mašine. Pri tome, pojave usled efekata rezultujućeg magnetskog zasićenja su izraženije kada je veći broj pobuđenih faza istog magnetskog polariteta. Pored toga, od interesa je istaći da na talasni oblik struje posmatrane faze, pored dominantnog uticaja susednih faza utiče i njena interakcija sa odgovarajućom ortogonalnom fazom. Taj efekat se manifestuje tokom značajnijeg istovremenog vođenja faza što se ima u ovom primeru, ali i tokom komutacije faza. Nesimetrija koja postoji između talasnih oblika faznih struja se preslikava i na odgovarajuće efektivne vrednosti koje su date u Tabeli 4.

Kao posledica rezultujućih efekata interakcije faza talasni oblik momenta je takav da su sva četiri impulsa momenta tokom električnog ciklusa različita saglasno Sl. 5. Imajući u vidu analizu uticaja magnetskog polariteta faza na fluksne obuhvate, kao i odgovarajuće talasne oblike faznih struja, u delu električnog ciklusa tokom kojeg se magnetiše faza A impuls momenta ima najveću trenutnu vrednost i to u trenutku demagntizacije faze D koja prethodi fazi A u sekvenci pobuđivanja. Nakon toga, slede vršne vrednosti impulsa momenta tokom magnetizacije faza D, B i C u trenucima demagnetizacije faza koje im prethode u sekvenci pobuđivanja a to su C, A i B respektivno.

Imajući u vidu rezultate sprovedene analize, uobičajena je praksa da se pobuđivanje 8/6 SRM-a vrši magnetskom sekvencom (S-N-S-N-N-S-N-S) kod koje su tri od četiri

TABELA 4
EFEKTIVNE VREDNOSTI FAZNIH STRUJA ZA RADNU TAČKU DEFINISANU U TABELI 3

I_A[A]	I_B[A]	I_C[A]	I_D[A]
4.5711	4.2149	4.2251	4.5572

kombinacije susednih faza motora različite, a jedna ista. Korišćenje ove sekvence pobuđivanja omogućava postizanje boljih performansi pogona sa 8/6 SRM-om u odnosu na analiziranu magnetsku sekvencu. Pri tome, u cilju postizanja najvećeg odnosa srednje vrednosti momenta i efektivne vrednosti struje ili ripla momenta, od suštinske je važnosti optimizacija kontrolnih parametara.

V. ZAKLJUČAK

U ovom radu je izvršena analiza uticaja magnetskog polariteta faza na statičke i tranzijentne karakteristike 8/6 SRM-a koristeći model koji se bazira na metodi konačnih elemenata u programskom paketu Ansys Electronics. Rezultati statičkih karakteristika pokazuju da je fluksni obuhvat posmatrane faze veći ako je ona u interakciji sa fazom različitog magnetskog polariteta u odnosu na nju. Pri tome, benefit od različitog magnetskog polariteta u odnosu na isti se osetnije javlja prilikom veće preklopljenosti statorskih i rotorskih polova. Sa druge strane, kada ne postoji njihovo preklapanje ili se ono smanjuje, tada su efekti međusobne interakcije faza manje izraženi. Odnos fluksnih obuhvata, čak i pri dubokom zasićenju, je približno isti u oba slučaja. Međutim, uvek je veći u slučaju različitog magnetskog polariteta. Pored toga, kada postoji veoma izraženo zasićenje jarma mašine, tada je usled magnetske nesimetrije faza veoma izražena nesimetrija faznih struja kao i impulsa u talasnom obliku momenta. Imajući u vidu gore navedeno, u cilju postizanja boljih performansi 8/6 SRM-a, faze motora se pobuđuju na način da su odgovarajući magnetski polariteti sukcesivnih faza motora dominantno različiti.

LITERATURA

[1] Z. Yang, F. Shang, I. P. Brown, and M. Krishnamurthy, “Comparative study of interior permanent magnet, induction, and switched reluctance motor drives for EV and HEV applications,” *IEEE Trans. Transp. Electrific.*, vol. 1, no. 3, pp. 245-254, Oct. 2015.

[2] R. Cardenas *et al.*, “Control of switched reluctance generator for variable-speed wind energy applications,” *IEEE Trans. Energy Convers*., 2005, vol. 20, no. 4, pp. 781-791, Dec. 2005.

[3] A. Chiba *et al*., ‘Development of a rare-earth-free SR motor with high torque density for hybrid vehicles,’ *IEEE Trans. Energy Convers.*, vol. 30, no. 1, pp. 175–182, March 2015.

[4] V. Valdivia *et al.*, “Behavioral modeling of a switched reluctance generator for aircraft power systems,” *IEEE Trans. Ind. Electron.*, vol. 61, no. 6, pp. 2690-2699, June 2014.

[5] A. K. Jain and N. Mohan, “Dynamic modeling, experimental characterization, and verification for SRM operation with simultaneous two-phase excitation,” *IEEE Trans. Ind. Electron.*, vol. 53, no. 4, pp. 1238–1249, Jun. 2006.

[6] P. Pillay, Y. Liu, W. Cai, and T. Sebastian, “Multiphase operation of switched reluctance motor drives,” in *Proc. 32nd IAS Annu. Meeting*, vol. 1, 1997, pp. 310–317.

[7] Qu, B.; Song, J.; Liang, T.; Zhang, H., "Mutual coupling and its effect on torque waveform of even number phase switched reluctance motor," Electrical Machines and Systems, International Conference on, pp.3405-3410, Oct. 2008.

[8] J. C. Moreira and T. A. Lipo, “Simulation of a four phase switched reluctance motor including the effects of mutual coupling,” *Elect. Mach. Power Syst.*, vol. 16, pp. 281–289, 1989.

[9] R. M. Davis and I. H. Al-Bahadly, “Experimental evaluation of mutual inductance in a switched reluctance motor,” in *4th IEEE Int. Conf. Power Electronics and Variable Speed Drives*, London, U.K., 1991, pp. 243–248.

[10] D. Panda, V. Ramanarayanan, “Mutual coupling and its effect on steady-state performance and position estimation of even and odd number phase switched reluctance motor drive,” *IEEE Transactions on Magnetics*, vol. 43, pp. 3445-3456, August 2007.

[11] Essah, D. N.; Sudhoff, S.D., "An improved analytical model for the switched reluctance motor," Energy Conversion, IEEE Transactions on , vol.18, no.3, pp.349,356, Sept. 2003.

[12] M. A. Preston and J. P. Lyons, “A switched reluctance motor model with mutual coupling and multi-phase excitation,” *IEEE Trans. Magn.*, vol. 27, no. 6, pp. 5243–5245, Nov. 1991.

[13] S. Cao and K. J. Tseng, “Evaluation of neighboring phase coupling effects of switched reluctance motor with dynamic modeling approach,” in *Proc. PIEMC 2000*, vol. 2, 2000, pp. 881–886.

[14] A. K. Jain and N. Mohan, “Dynamic modeling, experimental characterization, and verification for SRM operation with simultaneous two-phase excitation,” *IEEE Trans. Ind. Electron.*, vol. 53, no. 4, pp. 1238–1249, Jun. 2006.

[15] R. Mikail, I. Husain, and M. Islam, Finite element based analytical model for controller development of switched reluctance machines, in Proceedings of the Energy Conversion Congress and Exposition (ESSE 2013), Denver, pp. 920-925, September 2013.

[16] D. S. Mihic, M. V. Terzic, and S. N. Vukosavic, “A New Nonlinear Analytical Model of the SRM With Included Multiphase Coupling,” *IEEE Trans. Energy Convers.*, 2017, 32, (4), pp. 1322-1334.

ABSTRACT

To improve performance, the operation of a switched reluctance motor (SRM) requires simultaneous excitation of multiple motor phases. Since the magnetic polarity of excited phases can be the same or different, the influence of magnetic polarity on the flux linkages of 8/6 SRM was analyzed in this paper in detail. The analysis is performed based on static characteristics obtained using a machine model in a Finite Element Method (FEM) software. Furthermore, considering that the magnetic symmetry between phases cannot be achieved in an 8/6 SRM when using an Asymmetrical Half-Bridge Converter (AHBC), the resulting phase interactions are highly pronounced and are studied by analyzing waveforms of phase currents and torque.

Analysis of influence magnetic phase interaction on the characteristics of 8/6 SRM

Dragan S. Mihić,
Mladen V. Terzić,
Žarko V. Koprivica,
Bogdan M. Brković

Skin effect implementation in parameterized winding function model of an induction motor

Aldin Kajević, Mario Mezzarobba, Alberto Tessarolo, *Senior Member, IEEE* and Gojko Joksimović, *Senior Member, IEEE*

***Abstract* — The paper develops a method for skin effect implementation in recently derived parameterized winding function model of cage rotor induction motor. In that model number of rotor bars is free parameter. For any different number of rotor bars, rotor slot dimensions are different in order to preserve the total rotor copper volume but the slot shape is preserved. By defining the function of slot shape and using multilayer approach, rotor bar resistance and slot reactance can be calculated for any actual rotor speed and any number of rotor bars. The results from the model are given for two different number of rotor bars.**

***Index Terms* — Cage rotor induction motor, Winding function, Parameterized winding function, Multilayer approach, Skin effect.**

I. INTRODUCTION

The skin effect is a well-known and well described phenomenon that occurs in all conductors through which alternating current flows. This effect leads to the redistribution of current across the conductor cross section, which has an effect similar to reducing the cross-section area of the conductor. Non-uniform distribution is more and more significant as AC current frequency grows. The skin effect is usually undesirable because it leads to an increase of Joule losses and thus to increased heating of the conductors.

This effect is especially interesting in cage induction motors because it may have a positive effect. As it is well known from the basic principles of operation of an induction motor, the highest frequency in the rotor bars occurs at the motor startup and therefore the greatest bar resistance occurs during the motor starting. This is desirable because it leads to an increase in the value of the starting torque. On the contrary, when the motor rotates at rated speed, the slip frequency is very small and the increase in resistance due to the skin effect is negligible. By other words, rotor bar current distribution is uniform in that case. On the other side, the skin effect leads to a decrease in rotor bar leakage inductance.

Quantitative measures of skin effect are correction factors for resistance and leakage inductance. The resistance correction factor is the ratio of Joule losses in two cases: AC and DC case. Similarly, the leakage inductance correction factor is the ratio of magnetic energy in two cases: AC and DC case.

In a recently developed parameterized winding function (PWF) model, [1], where number of rotor bars and its skewing angle appears as free parameters, skin effect was neglected. The reason for that was in the fact that mentioned model is up to now predominantly used for analysis of rotor slot harmonics appearance in stator current spectrum, in steady state conditions. By other words, PWF model was up to now primarily used for purpose of finding the best possible solution for number of rotor bars from the electromagnetic torque ripple point of view in steady state conditions, [2], [3], [4], [5], [6], [7]. In order to define a more comprehensive model that would be applicable to transient conditions too, this effect should certainly be considered, and this is the real and basic motivation of this paper. The skin effect in this paper is considered by multilayer approach. In both the analyzed cases, rotors with Q_r=22 and Q_r=28 bars, skewed rotor bars are analyzed where the angle of skewing is equal to one stator slot pitch, $\gamma=2\pi/Q_s$, where Q_s is number of stator slots, [7].

II. BASICS OF PWF MODEL

PWF model enables changes of number of rotor bars in cage rotor induction motor by preserving its rated power, main machine dimensions and stator winding design. Additionally, the model enables changes in skewing angle of rotor bars. Both parameters are freely selectable variables in that model, [1].

The main idea is preserving the volume of rotor cooper (more exactly, rotor aluminum) for any different number of rotor bars that have a real sense. Therefore, when changing the number of rotor bars, Q_r, where nothing changes on the stator side, the cross-sectional area of the bar and end-ring must be changed in order to preserve the rated power of the motor, Fig 1. In order for the motor to develop the same power, the following equations must be satisfied, [1]:

$$Q_r R_b I_b^2 = Q_{r\_new} R_{b\_new} I_{b\_new}^2 \tag{1}$$

$$Q_r R_b \left(\xi \frac{2k_{w1} w_1 m_1}{Q_r} I_{1rated} \right)^2 = Q_{r\_new} R_{b\_new} \left(\xi \frac{2k_{w1} w_1 m_1}{Q_{r\_new}} I_{1rated} \right)^2 \tag{2}$$

Gojko Joksimović and Aldin Kajević are with Faculty of Electrical Engineering, University of Montenegro, Cetinjski put b.b., 81000 Podgorica, Montenegro (Gojko.Joksimovic@ucg.ac.me, Aldin.Kajevic@ucg.ac.me).
Alberto Tessarolo and Mario Mezzarobba are with Department of Engineering and Architecture, University of Trieste, Via Alfonso Valerio, 10-34127, Trieste, Italy (atessarolo@units.it, mmezzarobba@units.it).

$$\xi \cong 0.8\cos\varphi_n + 0.2 \tag{3}$$

from which new rotor bar resistance and cross section area result as follows:

$$R_{b\_new} = \frac{Q_{r\_new}}{Q_r} R_b \tag{4}$$

$$A_{b\_new} = \frac{Q_{r\_new}}{Q_r} A_b \tag{5}$$

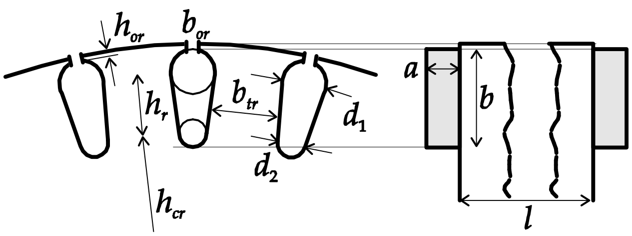


Fig. 1. Shape and dimensions of the rotor bar and the end-ring

In this way, rotor slot dimensions are determined for each new number of bars:

$$d_1 = \frac{\pi\left(D_{er} - 2h_{or}\right) - Q_r b_{tr}}{\pi + Q_r} \tag{6}$$

$$d_2 = \sqrt{\frac{8CA_b - \left(C\pi + 8\right)d_1^2}{C\pi - 8}} \tag{7}$$

$$h_r = \frac{d_1 - d_2}{2\tan\left(\pi/Q_r\right)} \tag{8}$$

$$b_{tr} \cong \frac{B_g \tau_r}{K_{Fe} B_{tr}} = \frac{B_g \pi D_r}{K_{Fe} B_{tr} Q_r} \tag{9}$$

$$C = 4\tan\left(\pi/Q_r\right) \tag{10}$$

III. MULTILAYER APPROACH

There are several ways in which skin effect i.e. correction factors for resistance and leakage inductance can be derived. For some simple shapes of slots there are analytical expressions while for more complicated shapes the bar is divided into several segments along the height of the bar - method known as the multilayer approach, [8]. The skin effect can also be considered using models based on the finite element method [9].

The multilayer approach method and a way of its implementation that is applicable in the PWF model is discussed here. The implementation of this method is usually time consuming and does not allow for a simple change in the dimensions and number of segments being analyzed. Here the previous problem was overcome by defining the function of the slot shape.

One of the possible ways to calculate skin effect in rotor bars of cage induction motor is to divide the bar into N layers, along the bar height. All layers are of the same height, Δh, but, in the general case, as a consequence of the shape of the rotor bar (rotor slot) and the position of the layer, the layers are of different widths, b_j, j=1, 2, 3, ..., N.

Each layer of the rotor bar is considered as a separate conductor of resistance R_j through which the current $\underline{I}_j$ flows. For the n-th layer, according to Faraday's law of electromagnetic induction, the following voltage equation can be written,

$$R_n \underline{I}_n - R_{n+1}\underline{I}_{n+1} = -js\omega_s \Delta\underline{\Phi}_n \tag{11}$$

where R_n, R_{n+1} and $\Delta\Phi_n$ are defined as follows:

$$R_n = \frac{1}{\sigma_{Al}} \frac{l}{b_n \Delta h} \tag{12}$$

$$R_{n+1} = \frac{1}{\sigma_{Al}} \frac{l}{b_{n+1} \Delta h} \tag{13}$$

$$\Delta\underline{\Phi}_n = \frac{\mu_0 l \Delta h}{b_n} \sum_{j=1}^{n} \underline{I}_j \tag{14}$$

Leakage inductance of the n-th layer is:

$$L_n = \frac{\mu_0 l \Delta h}{b_n} \tag{15}$$

Taking into account (14) from (11) we now obtain:

$$\underline{I}_{n+1} = \frac{R_n}{R_{n+1}} \underline{I}_n + j\frac{s\omega_s L_n}{R_{n+1}} \sum_{j=1}^{n} \underline{I}_j \tag{16}$$

Using the previous expression and knowing the current in the first layer, currents in all other layers can be obtained. To the current $\underline{I}_1$ one can assign an arbitrary value as it has no effect on the value of the final resistance and leakage inductance coefficients. These coefficients depend only on the slip frequency, bar shape and its dimensions.

After assigning a value to the current in the first layer, the currents in all other layers can be determined by iterative application of expression (16). When the currents in all layers are known, the value of total Joule losses in the bar can be determined taking into account the distribution of currents in the layers determined in the aforementioned manner, which is a consequence of the alternating current in the bar:

$$P_{AC} = \sum_{j=1}^{N} R_j \left|\underline{I}_j\right|^2 \tag{17}$$

Magnetic energy stored in the rotor slot can be defined in a similar way by taking into account the distribution of currents in the layers:

$$W_{AC} = \frac{1}{2}\sum_{j=1}^{N} L_j \left|\sum_{i=1}^{j} \underline{I}_i\right|^2 \tag{18}$$

If the total rms value of the bar current is defined as follows,

$$I_b = \left|\sum_{j=1}^{N} \underline{I}_j\right| \tag{19}$$

then we can calculate losses and magnetic energy in the bar when a direct current of the same intensity is supposed to flow through it,

$$P_{DC} = \sum_{j=1}^{N} R_j I_{jDC}^2 \tag{20}$$

where:

$$I_{jDC} = \frac{I_b}{A_b} b_j \Delta h \tag{21}$$

$$W_{DC} = \frac{1}{2}\sum_{j=1}^{N} L_j \left(\sum_{i=1}^{j} I_{iDC}\right)^2 \tag{22}$$

Correction coefficients for the resistance and inductance of the rotor bar are finally:

$$K_R = \frac{P_{AC}}{P_{DC}} = \frac{\sum_{j=1}^{N} R_j \left|\underline{I}_j\right|^2}{\sum_{j=1}^{N} R_j I_{jDC}^2} \tag{23}$$

$$K_L = \frac{W_{AC}}{W_{DC}} = \frac{\sum_{j=1}^{N} L_j \left|\sum_{i=1}^{j} \underline{I}_i\right|^2}{\sum_{j=1}^{N} L_j \left(\sum_{i=1}^{j} I_{iDC}\right)^2} \tag{24}$$

IV. Method implementation

The first step in applying the method is to determine the dimensions of each layer i.e. to determine the height Δh and the widths of the layers, b_j. One way to achieve this is by defining the slot shape function. The slot shape function for the slot shape used in this model is obtained using the analytical expressions for a circle and a line through two points. The coordinates of the points and the dimensions of the diameters that appear in the analytical expressions are recalculated for each new number of bars. A similar procedure can be applied to all slot shapes consisting of simple geometric shapes for which there are analytical expressions, which is the most common case. For the shape of the slot used in this model, the following function can be defined,

$$f(x) = \begin{cases} \sqrt{x(d_2 - x)}; & 0 \le x \le 0.5d_2 \\ \dfrac{(d_1 - d_2)(2x - d_2) + 2d_2 h_r}{4h_r}; & 0.5d_2 < x \le 0.5d_2 + h_r \\ 0.5\sqrt{d_1^2 - (2x - d_2 - 2h_r)^2}; & 0.5d_2 + h_r < x \le 0.5(d_1 + d_2) + h_r \end{cases} \tag{25}$$

where x is the position of the layer measured from the bottom of the slot upwards, Fig. 2.

If the height of the layers is Δh, same for all of them, and x is defined as follows,

$$\Delta h = \frac{0.5(d_1 + d_2) + h_r}{N} \tag{26}$$

$$x = 0,\ \Delta h,\ 2\Delta h,\ \dots N\Delta h \tag{27}$$

then the widths of the layers b_j can be obtained by applying (25) for different values of x:

$$b_j = 2f(x_j) \tag{28}$$

Once the dimensions of all layers have been defined, the procedure for determining the correction coefficients can be applied. For the shape of the bar used in this model, a graph of the dependency of the correction factors upon slip has been obtained, as shown in Fig. 3.

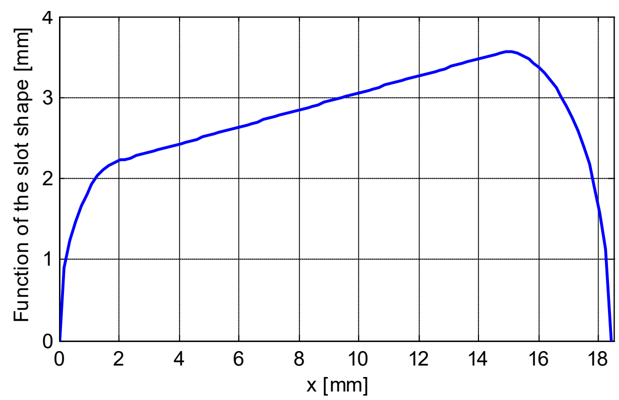


Fig. 2. Rotor slot shape function. Number of rotor bars, Q_r=30.

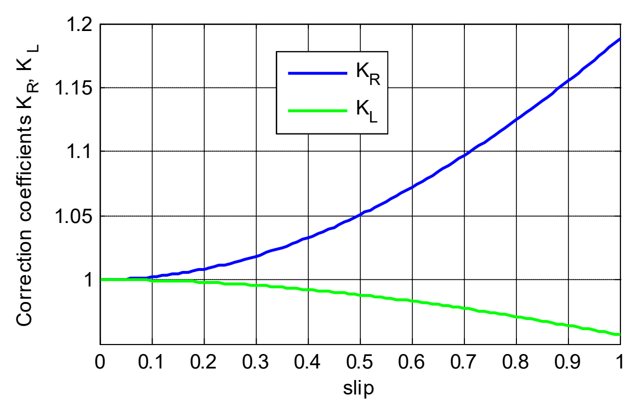


Fig. 3. Correction coefficients as a function of slip. Number of rotor bars Q_r=30.

V. Results

Using the previous expressions, correction coefficients can be determined for different bar numbers. The procedure for determining the correction factors can be incorporated into the parameterized dynamic model based on the winding function theory (PWF model) [4]. Figures below show comparisons of the results obtained from this model without and with the skin effect taken into account, for two different numbers of rotor bars, Q_r=22 and Q_r=28. In both analyzed cases higher electromagnetic torque is an obvious consequence and therefore, the steady-state rotor speed is reached faster.

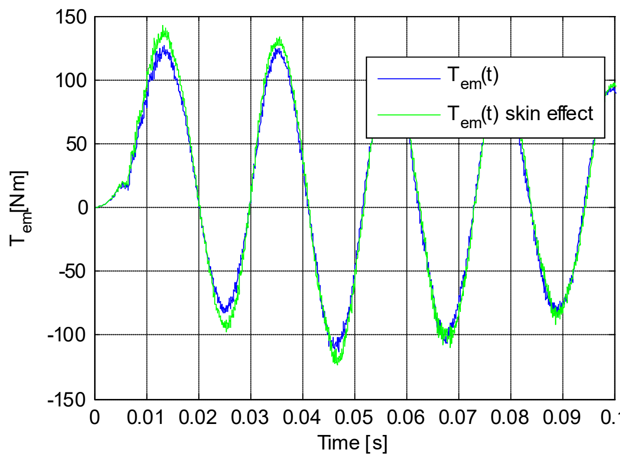


Fig. 4. Electromagnetic torque during the no-load speed-up of the motor, Q_r=22

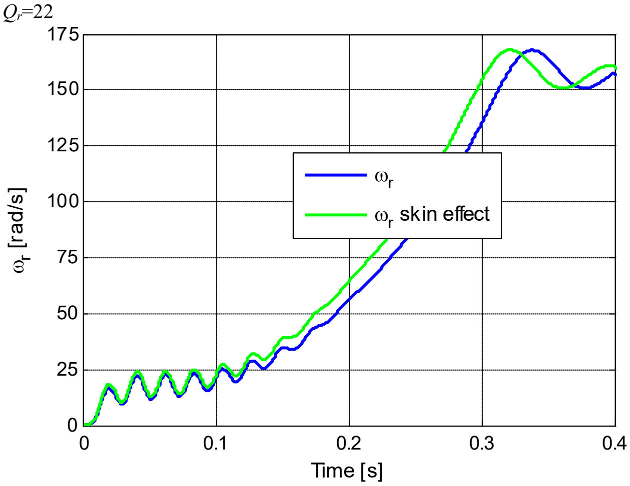


Fig.5. Rotor speed during the no-load, speed-up of the motor, Q_r=22

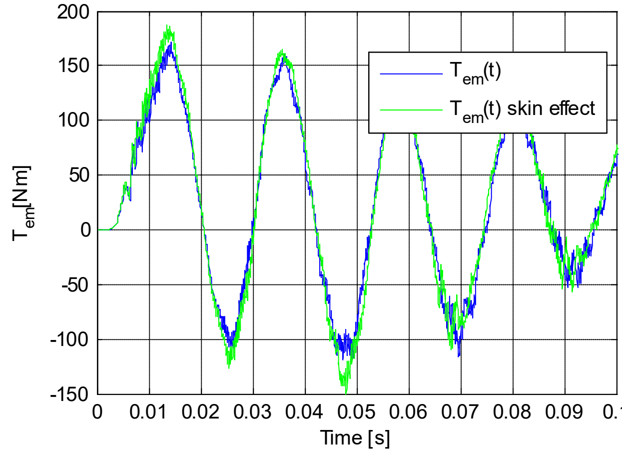


Fig. 6. Electromagnetic torque during the no-load speed-up of the motor, Q_r=28

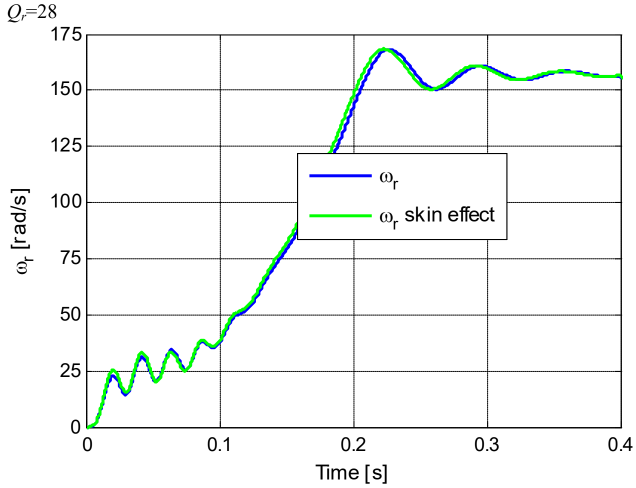


Fig. 7. Rotor speed during the no-load speed-up of the motor, Q_r=28

VI. Conclusion

Multilayer approach method and a modified way of its implementation that allows a simple change in the dimensions of the slot as well as the number of segments used in the analysis are presented in this paper. This was realized by defining the function of the slot shape. The method was incorporated in parameterized dynamic model based on the winding function theory. The results from the model are shown and illustrated. The influence of the skin effect on the motor starting torque can be clearly seen from the presented results for two different number of rotor bars.

This model does not consider saturation of ferromagnetic material. Further improvements to the model are planned, one of which is to take into account saturation. It is also planned to compare the results with the results obtained using the FEM model, which can consider the saturation effects.

APPENDIX

TABLE I
MOTOR RATED VALUES, MOTOR AND ROTOR SLOT GEOMETRICAL PARAMETERS

P_r [kW]	11	Q_s	36
U_{LL} [V]	400	Q_r	30
f [Hz]	50	y/τ	7/9
I_r [V]	17.6	q	3
$\cos\varphi_r$	0.83	W_1	108
η_r	0.91	D_{is} [mm]	145.724
p	2	L [mm]	171.677
R_s [Ω]	0.294	g [mm]	0.397
$L_{\sigma s}$ [mH]	2.919	d_1 [mm]	7.132
R_b [μΩ]	64.49	d_2 [mm]	4.480
R_{er} [μΩ]	1.545	h_r [mm]	12.615
L_b [nH]	398.58	A_b [mm^2]	101.092

ACKNOWLEDGMENT

This research was conducted on Faculty of Electrical Engineering, University of Montenegro and was financed by the Ministry of Science of Montenegro through the research project “Induction motor efficiency improvement through optimal electromagnetic design solutions - IMEI”. Partner on the project is Department of Engineering and Architecture, University of Trieste, Italy.

REFERENCES

[1] G. Joksimović, “Dynamic model of cage induction motor with number of rotor bars as parameter,” *J. Eng.*, vol. 2017, no. 6, pp. 205–211, Jun. 2017.
[2] G. Joksimović, J. I. Melecio, P. M. Tuohy, S. Djurović, „Towards the optimal ‘slot combination’ for steady-state torque ripple minimization: an eight-pole cage rotor induction motor case study”, Electrical Engineering, Springer, vol. 102, Issue 1, pp. 293-308, 2020.
[3] G. Joksimović, M. Mezzarobba, A. Tessarolo, E. Levi, „Optimal Selection of Rotor Bar Number in Multiphase Cage Induction Motors”, IEEE Access, vol. 8, pp. 135558-135568, 2020.
[4] G. Joksimović, E. Levi, A. Kajević, M. Mezzarobba, A. Tessarolo, „Optimal Selection of Rotor Bar Number for Minimizing Torque and Current Pulsations Due to Rotor Slot Harmonics in Three-phase Cage Induction Motors”, IEEE Access, vol. 8, pp. 228572-228585, 2020.
[5] G. Joksimović, A. Kajević, M. Mezzarobba, A. Tesarolo, “Optimal rotor bars number in four pole cage induction motor with 36 stator slots - part I: numerical modeling“, ICEM 2020, Gothenburg, Sweden, 2020.
[6] G. Joksimović, A. Kajević, M. Mezzarobba, A. Tesarolo, “Optimal rotor bars number in four pole cage induction motor with 36 stator slots - part II: results“, ICEM 2020, Gothenburg, Sweden, 2020.
[7] G. Joksimović, A. Kajević, S. Mujović, T. Dlabač, V. Ambrožič, A. Tesarolo, “Rotor bars skewing impact on electromagnetic pulsations in cage induction motor“, IcEtran 2019, Srebrno jezero, Srbija, 2019.
[8] I. Boldea, S. A. Nasar, The induction machine handbook, CRC Press, 2010.
[9] S. Williamson and M. J. Robinson, “Calculation of cage induction motor equivalent circuit parameters using finite elements,” IEE Proc. B Electr. Power Appl., vol. 138, no. 5, pp. 264–276, 1991.

Operation Analysis and Determination of Virtual Synchronous Machine Model Parameters

Nikola Krstić, Milutin Petronijević, Filip Filipović

Abstract **–This paper presents the concept, model and simulation of virtual synchronous machine (VSM) operation and proposes a methodology for its parameter values selection. VSM was synchronized to the grid and its responses to given references of active and reactive powers were obtained, for different values of VSM parameters. Responses of grid-feeding and grid-supporting model of VSM, were considered. Special attention is paid to the formation of control loops for active and reactive power as basic structures in the VSM operation. Based on these control loops, values of VSM parameters are determined by pole adjustment method, which was the basic task of this paper. VSM model and the model of used power grid were made in Matlab/Simulink, in which results presented at the end of the paper, were generated.**

Keywords – **virtual synchronous machine, power converter, control loop, emulation.**

I. Introduction

Increased distributed generation, presence of renewable energy sources, as well as the emergence of microgrids in the power system, have led to new ideas and concepts to appear [1]. Many of these concepts include the use of power electronics to increase the flexibility and range of power grids, and one of them is virtual synchronous machine (VSM). This concept arose from the need to bring the operation of three-phase power converters closer to the operation of synchronous machines [2], primarily by introducing virtual inertia. Doing this, the control operation of energy converters has not changed significantly, and it is still based on droop control (which is inherited from synchronous machines), but now a certain inertia is included in it. Addition of inertia, which emulates moment of inertia of the rotor of synchronous machine and the inertia of inductance in excitation circuit, increases stability of the grid and its resistance to various types of disturbances and power imbalances [3]. This is especially important in grids where large proportion of generated power comes from power sources connected via power converters [4].

In this paper, structure and concept of operation of VSM model [3],[5],[6] are presented first, where the equations of mathematical model of synchronous machine, that is going to be emulated with VSM, are written. After that, the methodology for determining the values of VSM parameters was presented. It is clear that values of VSM model parameters will largely determine its response [7] to the given references, which is why their proper selection, in order to meet a particular requirement, is the main task of this paper. In order to make this possible, VSM model was decomposed into two basic structures, one includes active power-frequency regulation, and the other reactive power-voltage regulation [5]. In this way, control loops for active and reactive power are formed in the paper, on the basis of which values of VSM parameters are determined, using pole adjustment method. Number of these parameters differs in the case of grid-feeding and grid-supporting model of VSM, that depends on the role which energy converter have in the grid [8]. Grid-feeding energy converter only meets power requirements regardless of the grid voltage and frequency, while grid-supporting one, in addition to that, supports regulation of voltage and frequency in the grid. That is way, in this paper, two additional coefficients in the grid-supporting model of VSM are used, one for grid voltage and the other for grid frequency control. Three basic parameters, common to both VSM models, are: moment of inertia, damping factor and ratio coefficient of reactive power to change of excitation flux [3]. In this paper, VSM is connected to the grid via LCL filter and the grid is modeled by real voltage generator.

Results presented at the end of the paper contain responses of certain quantities, for different values of parameters [9] in VSM model for grid-feeding and grid-supporting structure.

Nikola Krstić is with the Department of Power Engineering, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: nikola.krstic@elfak.ni.ac.rs.

Milutin Petronijević is with the Department of Power Engineering, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: milutin.petronijevic@elfak.ni.ac.rs.

Filip Filipović is with the Department of Power Engineering, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: filip.filipovic@elfak.ni.ac.rs.

II. VSM Model

Basic idea behind the concept of VSM is to create an emulator which will ensure that three-phase power converter behaves like a synchronous machine towards the grid. That is, it is necessary to equalize voltages at the end of the power converter with electromotive forces that would occur in the operation of a synchronous machine. Since the electromotive forces in synchronous machine are determined by the first derivative of excitation flux, it is clear that VSM model must take into account relations for excitation flux and angular frequency. In this paper, mathematical model of synchronous machine with cylindrical rotor and without losses, is used to form VSM model. Excitation flux through the phase windings of synchronous machine is given by equations:

$$\psi_{fA} = \psi_{fm}\cos(\omega t) \quad (1)$$

$$\psi_{fB} = \psi_{fm}\cos(\omega t - \tfrac{2\pi}{3}) \quad (2)$$

$$\psi_{fC} = \psi_{fm}\cos(\omega t + \tfrac{2\pi}{3}) \quad (3)$$

Based on excitation flux, electromotive forces in phase windings are determined according to the vector equation:

$$\vec{e}_f = -\frac{d\vec{\psi}_f}{dt} \quad (4)$$

Based on (4), electromotive forces in phase windings can be obtain using equations:

$$e_{fA} = \omega\psi_{fm}\sin(\omega t) - \frac{d\psi_{fm}}{dt}\cos(\omega t) \quad (5)$$

$$e_{fB} = \omega\psi_{fm}\sin(\omega t - \tfrac{2\pi}{3}) - \frac{d\psi_{fm}}{dt}\cos\left(\omega t - \tfrac{2\pi}{3}\right) \quad (6)$$

$$e_{fC} = \omega\psi_{fm}\sin(\omega t + \tfrac{2\pi}{3}) - \frac{d\psi_{fm}}{dt}\cos\left(\omega t + \tfrac{2\pi}{3}\right) \quad (7)$$

Angular frequency of VSM appearing in (1)-(7) is found by motion equation of rotor of VSM:

$$\frac{d\omega}{dt} = \frac{1}{J}\left(\frac{P_{set}}{\omega} - \frac{P}{\omega} - D_p(\omega - \omega_G)\right) \quad (8)$$

Where: J −moment of inertia of rotor of VSM, ω −angular frequency of VSM, ω_G −angular frequency of the grid, D_p −damping factor, P_{set} −active power reference (generated mechanical power), P −active power injected into the grid by VSM (active power of consumption).

It must be noted that in (8) generated mechanical power of VSM is identified as active power reference and generated active power is identified as active power injected into the grid. This identifications can be done because model of VSM emulates synchronous machine without losses and active power losses in LCL filter are negligible. In addition to the angular frequency, it is necessery to determine the intensity of excitation flux in VSM. This is done on the basis of excitation regulation in synchronous machine, using reactive power eror:

$$\frac{d\psi_f}{dt} = D_q(Q_{set} - Q) \quad (9)$$

Where: ψ_f −excitation flux in phase windings of VSM, D_q −ratio coefficient of reactive power to change of excitation flux, Q_{set} −reactive power reference, Q −reactive power injected into the grid by VSM.

Based on (5)-(9), grid-feeding model of VSM is formed in Matlab/Simulink. In the case of grid-supporting model, additional links for grid voltage and frequency regulation must be considered. In this paper, grid is modeled with real voltage generator whose equivalent impedance has big *R/X* ratio, that is usually the case in low voltage networks. This would mean that the grid voltage regulation in the point of connection of VSM to the grid is most efficiently achieved using injected active power. On the other hand, the best way to regulate grid frequency in the point of connection is injecting reactive power. This is based on the voltage drop formula between node *i* and node *j* connected with branch containing active resistance *R* and reactance *X*, when active power P_{ij} and reactive power Q_{ij} flows from node *i* to node *j*.

$$U_i - U_j = \frac{P_{ij}R + Q_{ij}X}{U_i} + j\frac{P_{ij}X - Q_{ij}R}{U_i} \quad (10)$$

Considering (10), in *R* dominant type of grids, voltage in connection point (U_i) is increased when injected active power (P_{ij}) is increased, and the phase ange (frequency) is increased when injected reactive power (Q_{ij}) is decresed [1],[8]. Considering this, for grid-supporting model, (8)-(9) become:

$$\frac{d\omega}{dt} = \frac{1}{J}\left(\frac{P_{set} - P + K_V(V_{ref} - V_G)}{\omega} - D_p(\omega - \omega_G)\right) \quad (11)$$

$$\frac{d\psi_f}{dt} = D_q\left(Q_{set} - Q - K_\omega\left(\omega_{ref} - \omega_G\right)\right) \quad (12)$$

Where: V_G −grid voltage, V_{ref} −grid voltage reference, K_ω −control coefficient of grid angular frequency, K_V −control coefficient of grid voltage, ω_{ref} −grid angular frequency reference.

Signals of obtained electromotive forces given by (5)-(7) are used to generate voltage at the end of the energy converter, what can be seen in Fig. 1, where grid-supporting model of VSM is presented, as more general one.

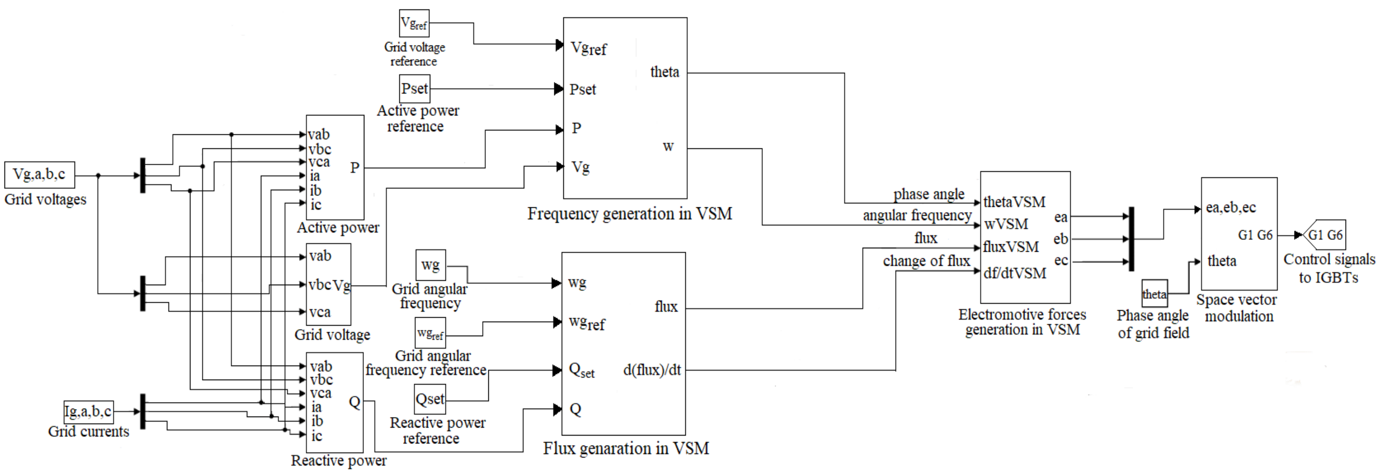


Fig. 1. Block diagram of grid-supporting model of VSM

III. Formation of Control Loops and VSM Parameter Determination

In order to determine the parameters in VSM model, it is necessary to form appropriate control loops. Due to the droop control, this can be most easily and efficiently done by splitting VSM model into two independent units, one related to active power and other to reactive power regulation. Based on these units, control loops for active and reactive power are formed, shown in Figs 2 and 3, respectively. During their formation, certain assumptions were adopted in order to achieve linearization and complete independence between the loops, and thus simplify the calculation of VSM parameters. This refers to the use of indicated values for the grid voltage and angular frequency of VSM instead of their actual values. Also, fixed value of electromotive force in the regulation of active power and fixed load angle in regulation of reactive power is used. Second member in expressions for electromotive force (5)-(7), is neglected.

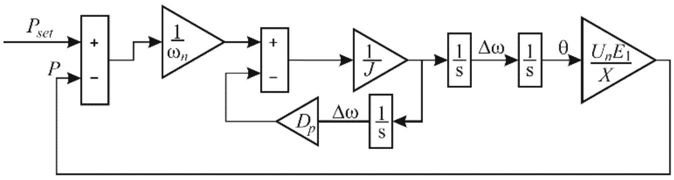


Fig. 2. Active power control loop

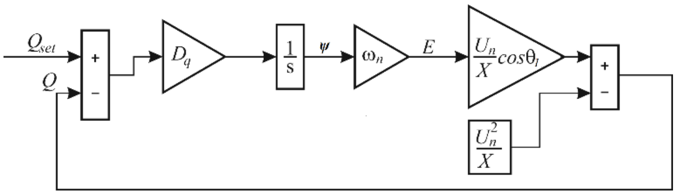


Fig. 3. Rective power control loop

Linearization was also achieved by replacing the sinusoidal function with its argument in expression for active power (13) and using fixed value for cosine function in expression for reactive power (14). These assumptions are justified if deviations of grid operating voltage and VSM frequency from their indicated values are small and if load angle does not take big changes. Also, product of flux and angular frequency of VSM has to be an order of magnitude greater than the change of VSM flux. All of this, to a large extent, should be met if values for active and reactive power are set in normal operating mode. Biger errors can occur by using inadequate fixed values for electromotive force and load angle of VSM in control loops. In order to partially overcome this, mentioned quantities were calculated on the basis of set values for active and reactive power.

In this paper, VSM emulates the operation of synchronous machine without losses and with cylindrical rotor, which active and reactive power can be determined by expressions:

$$P = \frac{U \cdot E}{X} sin\theta \tag{13}$$

$$Q = \frac{U \cdot E}{X} cos\theta - \frac{U^2}{X} \tag{14}$$

Where: P −active power that VSM injects into the grid, Q −reactive power that VSM injects into the grid, U −grid voltage, E −electromotive force of VSM, θ −load angle, X −equivalent reactance between electromotive force and voltage.

Taking into account mentioned assumptions while forming control loops for active and reactive power, (13) and (14) become:

$$P = \frac{U_n \cdot E_1}{X} \theta \tag{15}$$

$$Q = \frac{U_n \cdot E}{X} cos\theta_1 - \frac{U_n^2}{X} \tag{16}$$

Where E_1 and θ_1 are fixed values for electromotive force and load ange, determined by set values of active and reactive power, using expressions:

$$E_1 = \frac{\frac{Q_{seti}+Q_{seti-1}}{2} \cdot X + U_n^2}{U_n} \tag{17}$$

$$\theta_1 = sin^{-1}\left(\frac{\frac{P_{seti}+P_{seti-1}}{2} \cdot X}{U_n \cdot E_1}\right) \tag{18}$$

Where indices i and $i-1$ in active and reactive power references mark current and previous reference, respectively.

Equivalent reactance between electromotive force of VSM and grid voltage, which appears in expressions for active and reactive power, is the equivalent reactance of LCL filter, and can be found as:

$$X = X_1 - \frac{X_2 \cdot X_C}{X_2 - X_C} \tag{19}$$

Where: X_1 −reactance of the inductance near the VSM ($X_1 = \omega L_1$), X_2 −reactance of the inductance near the grid ($X_2 = \omega L_2$), X_C −reactance of the capacitor ($X_C = 1/C\omega$). Delay due to the change of currents in the network branches is neglected in control loops, because it is much less than what is needed to change flux or frequency and practically does not affect the dinamics of the system. Based on the control loop for active power, transfer function between the set and actual active power injected into the grid by VSM, can be found as:

$$\frac{P}{P_{set}} = F_p(s) = \frac{\frac{U_nE_1}{X\omega_nJ}}{S^2 + \frac{D_p}{J}S + \frac{U_nE_1}{X\omega_nJ}} \tag{20}$$

In this paper, pole adjustment method is used, which means that VSM parameters are calculated from desired values of characteristic parameters that determine the response of second order circuit. General expression for second order transfer function is:

$$F_s(s) = \frac{\omega_C^2}{s^2 + 2\xi\omega_C s + \omega_C^2} \tag{21}$$

Where ω_C is the natural frequency (bandwidth of second order circuit) and ξ is the damping factor of second order circuit. Comparing (20) and (21), moment of inertia and damping factor of VSM can be determined based on the desired damping factor and bandwidth of second order circuit:

$$J = \frac{U_n \cdot E_1}{X \cdot \omega_n \cdot \omega_C^2} \tag{22}$$

$$D_p = 2 \cdot \xi \cdot \omega_c \cdot J \quad (23)$$

Same analogy can be used for reactive power control loop, whose transfer function is given by expression:

$$\frac{Q}{Q_{set}} = F_q(s) = \frac{1}{1+\frac{X}{D_q \omega_n U_n cos\theta_1}s} \quad (24)$$

General form of the first order transfer function is:

$$F_s(s) = \frac{1}{1+sT} \quad (25)$$

Where T is the time constant of the first order circuit. Based on the desired value of time constant, ratio coefficient D_q can be found as:

$$D_q = \frac{X}{T \cdot \omega_n \cdot U_n \cdot cos\theta_1} \quad (26)$$

Now, values are determined for all required parameters of grid-feeding model of VSM. Values for an additional two parameters in grid-supporting model of VSM are determined based on the desired error between set and actual values of injected powers, during the deviation of grid voltage and frequency from the reference values, in steady state. This is shown in (27) and (28):

$$K_V = -\frac{P-P_{ref}}{V-V_{ref}} \quad (27)$$

$$K_\omega = \frac{Q-Q_{ref}}{\omega-\omega_{ref}} \quad (28)$$

Expressions (27) and (28) are obtain using (11) and (12) for steady state. Determined in this way, control coefficients have positive values. It must be taken into account that (27) and (28) are valid only in active resistance dominant grids, like one used in this paper.

IV. Presentation and Analysis of Results

In this section, results related to the response of active and reactive power of VSM for different values of its parameters are presented and analyzed. Used active and reactive power references and the rating of VSM (40kVA) are such that they do not cause major voltage deviations in the formed low voltage grid model. All results were obtained for the case where VSM, powered by storage system DC voltage, is connected to the low voltage grid, modeled by real voltage generator, via LCL filter. This is shown in the form of Matlab/Simulink block in Fig. 4. For the *50 Hz* frequency grid resistance and grid reactance used in the model are: $R_g=0.1126\Omega$ and $X_g=0.0233\Omega$. Voltage generator in grid modeling is fixed 400V, 50Hz voltage source. Reactances of LCL filter are: $X_1=0.6283\Omega$, $X_2=0.1571\Omega$ and capacitor reactance $X_C=27.6791\Omega$ (equivalent reactance is $X=0.785\Omega$).

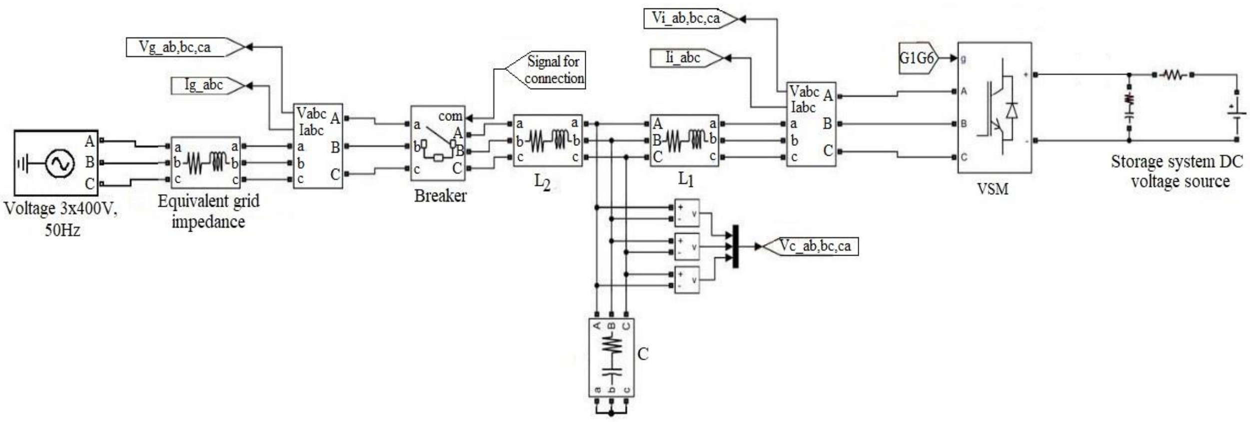


Fig. 4. Block model of power grid to which VSM, powered by storage system DC voltage, is connected via LCL filter

Next six figures show how efficient is the response of active and reactive power of grid-feeding model of VSM and to what extent it matches the desired (expected) responses. Desired response of active and reactive power is determined by chosen values of characteristic parameters ω_C, ξ and T. Based on those three characteristic parameters, using presented methodology, necessary values of parameters in grid-feeding model of VSM are determined, for three different cases, shown in Table I.

Table I
Characteristic parameters of the desired response and calculated parameters for grid-feeding model of VSM

Case	Natural frequency ω_C (rad/s)	Damping factor ξ	Time constant T (s)
1	10	0.707	0.15
2	7	1	0.15
3	14	0.5	0.2
Case	Moment of inertia J (kgm^2)	Damping coefficient $D_p(Nms/rad)$	Ratio coefficient $D_q(Wb/sVAr)$
1	6.4458	91.1441	$4.247 \cdot 10^{-5}$
2	13.1548	184.1667	$4.2454 \cdot 10^{-5}$
3	3.2887	46.0417	$3.184 \cdot 10^{-5}$

Figs 5 and 6 show reference, desired and actual response of active and reactive power for the case 1.

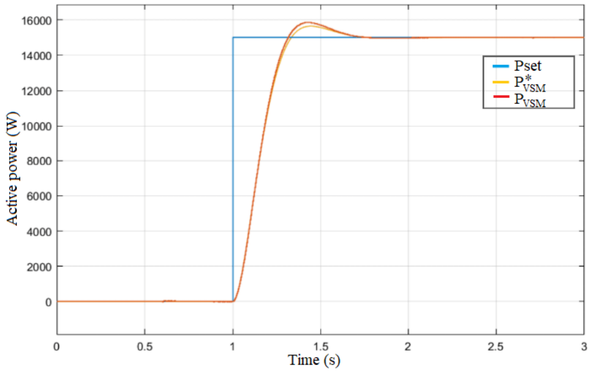


Fig. 5. Reference, desired and actual response of active power in grid-feeding model, case 1

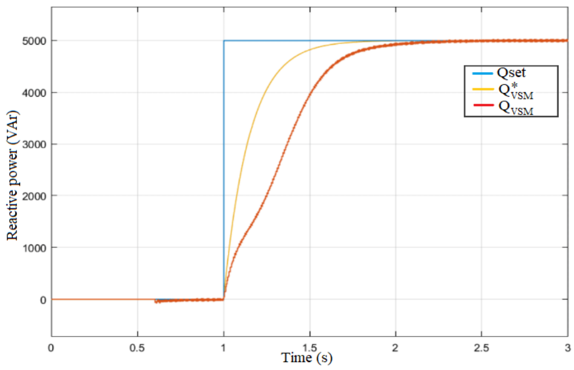


Fig. 6. Reference, desired and actual response of reactive power in grid-feeding model, case 1

Figs 7 and 8 show reference, desired and actual response of active and reactive power for the case 2.

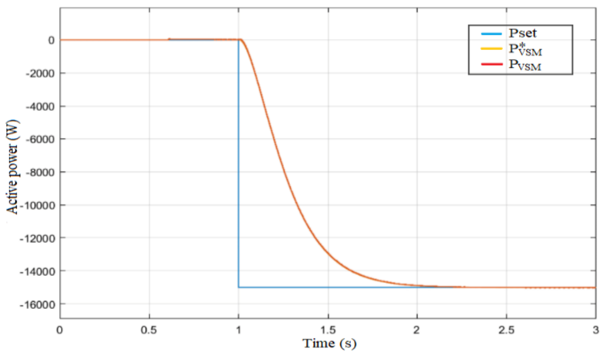


Fig. 7. Reference, desired and actual response of active power in grid-feeding model, case 2

Based on the Figs (5)-(10) it can be concluded that VSM operates efficiently and achieves desired references of active and reactive powers (positive or negative) with high precision. Also, great match between desired (expected) and actual active power response can be seen. In the case of reactive power, matching is not so great when active and reactive power have the same direction, in other cases it is quite good considering the simplicity of the methodology. This can be explained by the fact that voltage and frequency variations, have more affect on the dynamics of reactive than on the active power response, considering (20) and (24). Also, voltage variations are the largest when both powers have the same direction, that is way the worst matching is in the fig. 6.

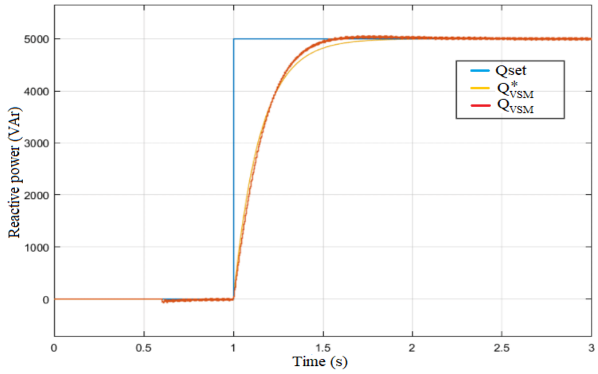


Fig. 8. Reference, desired and actual response of reactive power in grid-feeding model, case 2

Figs 9 and 10 show reference, desired and actual response of active and reactive power for the case 3.

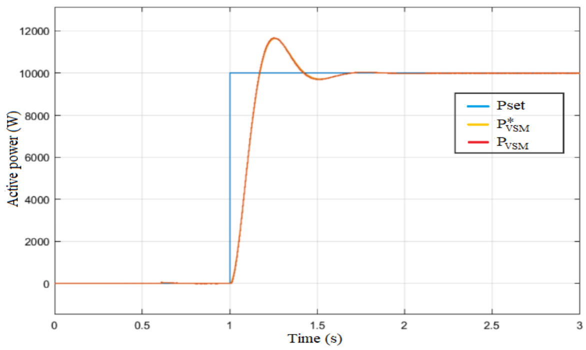


Fig. 9. Reference, desired and actual response of active power in grid-feeding model, case 3

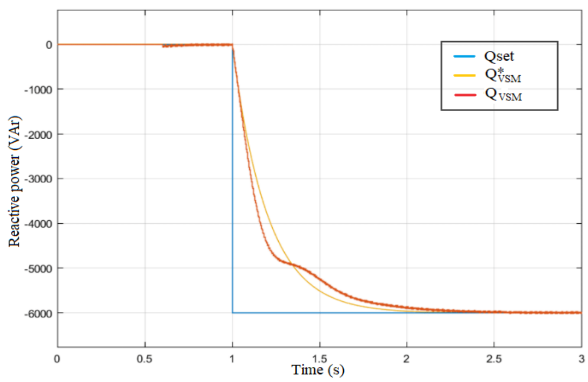


Fig. 10. Reference, desired and actual response of reactive power in grid-feeding model, case 3

In the grid-supporting model, references and actual values of injected powers will not match when there is a deviation in values of grid voltage and frequency compared to their referent values. This is shown in Fig. 11 for reactive and in Fig. 12 for active power.

TABLE II
CHARACTERISTIC PARAMETERS OF THE DESIRED RESPONSE AND CALCULATED PARAMETERS FOR GRID-SUPPORTING MODEL OF VSM

Natural frequency ω_C (rad/s)	Damping factor ξ	Time constant T (s)	Control coefficient $K_V(W/V)$
10	0.707	0.15	10 000
Control coefficient $K_\omega(VArs/rad)$	Moment of inertia $J\,(kgm^2)$	Damping coefficient $D_p(Nms/rad)$	Ratio coefficient $D_q(Wb/sVAr)$
30 000	6.3663	90.0189	$4.24\cdot 10^{-5}$

Table II contains values of characteristic parameters of the desired response, based on which values of parameters in grid-supporting model of VSM are calculated and also shown in Table II. Figs 11 and 12, are obtained for the case where parameters of grid-supporting model of VSM have values from Table II, and where grid voltage and frequency are below their referent values. High values for control coefficients are used due to fixed voltage source and low network impedance, which do not allow more significant changes in grid voltage and especially in grid angular frequency. There is a noticeable difference in both active and reactive power response shown in Figs 11 and 12, compared with those in Figs 5 and 6, which had the same values for characteristic parameters. This changed dynamics in grid-supporting model is the result of grid voltage and angular frequency errors used as inputs in active and reactive power control loops in which they are not considered in the first place. This means that presented methodology should be primarily used in grid-feeding model of VSM.

Considering the results for grid-feeding VSM model, it can be concluded that grid parameters have negative effect on active and reactive power response and decrease the accuracy of the proposed methodology by changing the voltage and frequency in the point of connection. This also applies to the grid-supporting model of VSM, but in that case voltage and frequency variations, caused by the grid impedance, in addition, enhance voltage and frequency support in the point of connection. Resistive or reactive nature of grid impedance determined which type of power should be used for voltage, and which for frequency support.

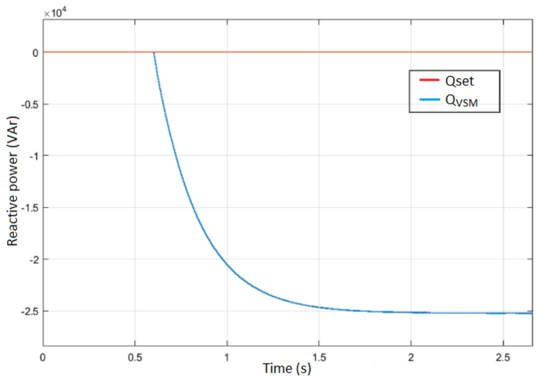


Fig. 11. Referent and actual value of reactive power in grid-supporting model

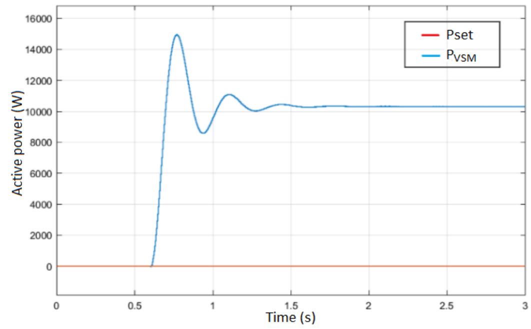


Fig. 12. Referent and actual value of active power in grid-supporting model

V. CONCLUSION

In this paper, model of VSM was made and methodology for determining the values of its parameters is presented and described in detail. Obtained results have shown that VSM is able to efficiently and accurately respond to different requirements of active and reactive powers. Proposed methodology for determining VSM parameter values is simple, and based on obtained results, quite accurate in achieving the desired response, if grid voltage and frequency deviations, due to power injection, are relatively small. Also, results have shown that if more precision in achieving the desired response is needed, proposed methodology should be primarily used in grid-feeding model of VSM. Taking into account the simplicity and efficiency of the presented methodology, its major practical value in real, more complex systems, could be reflected in determining the initial, first step, values of VSM parameters in order to achieve the desired response, which could then be corrected by additional steps, considering the characteristics of a particular system.

REFERENCES

[1] Q. Zhong, T. Hornik, *Control of Power Inverters in Renewable Energy and Smart Grid Integration*, John Wiley & Sons Ltd, Chichester, United Kingdom, 2013, pp. 251-330.

[2] M. Abuagreb, B. Ajao, H. Herbert, B. Johnson, "Evaluation of Virtual Synchronous Generator Compared to Synchronous Generator", IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), pp. 1-5, February 2020.

[3] J. M. Ramirez, E. T. Montalvo, C. I. Nuno, "Modeling, synchronization, and implementation of the virtual synchronous generator: a study of its reactive power handling", Electrical Engineering, pp. 1605-1619, March 2020.

[4] S. Vukosavić, *Grid-Side Converters Control and Design*, Springer International Publishing, 2018, pp. 211-250.

[5] H. Wu, X. Ruan, D. Yang, X. Chen, W. Zhao, Z. Lv, Q. Zhong, "Small-Signal Modeling and Parameters Design for Virtual Synchronous Generators", IEEE Transactions on Industrial Electronics, vol. 63, no.7, pp. 4292-4303, July 2016.

[6] Y. Hirase, K. Abe, K. Sugimoto, Y. Shindo, "A Grid Connected Inverter with Virtual Synchronous Generator Model of Algebraic Type", IEEJ Transactions on Power and Energy, vol. 132, no.4, pp. 371-380, April 2012.

[7] R. Aouini, K. Ben Kilani, B. Marinescu, M. Elleuch "Virtual Synchronous Generator Dynamic Performances", Internacional Conference on Electrical Science and Technologies in Magreb (CISTEM), pp. 1-6, November 2014.

[8] M. Castilla, L. Garcia de Vicuna, J. Miret, "Control of Power Converters in AC Microgrids", *Microgrids Desing and Implementation*, Springer Internacional Publishing, pp. 139-170, 2019.

[9] X. Yan, S. Mohamed, D. Li, A. Gadalla, "Parallel Operation of Virtual Synchronous Generators and Synchronous Generators in a Mycrogrid", Journal of Engineering, vol. 2019, pp. 2635-2642, March 2019.

Design of LLC Resonant Tank in a Low Power DC/DC Power Converter

Katarina Obradović, Emilija Lukić, Jovana Plavšić and Aleksandar Milić

***Abstract*— Photovoltaic power conversion prosperity has put a spotlight on the resonant DC/DC converters. Namely, improved power density and lessened power losses can be achieved due to its soft switching feature. Step by step design procedure of the LLC resonant tank is proposed in this paper. Analysis of the parameters of the tank, capacitor selection and a detailed inductor design are demonstrated thoroughly. Finally, experimental results of the developed 1000 W DC/DC converter with suggested LLC resonant configuration are presented.**

***Index Terms*— DC/DC converter, LLC resonant tank, high frequency, capacitor selection, inductor design**

I. Introduction

Increased interest in the use of renewable energy sources has inspired the development of new solutions in power electronics. Its unprecedented expansion has resulted in continued advances regarding power conversion's efficiency, safety, as well as the price.

Since photovoltaic (PV) systems represent DC power source, in order to successfully connect them to AC power grid, the use of switching power converters is inevitable. With aspirations to achieve better efficiency of power conversion, Maximum Power Point Tracking (MPPT) algorithm needs to be implemented in a DC/DC stage of the device. Possible topologies include non-isolated converters such as boost, buck, buck-boost or Ćuk topology presented in [1] - [4] and isolated configurations like flyback, push-pull, and resonant converters [5] - [8]. Furthermore, in a MPPT stage soft switching can be achieved and high frequency employed [9].

High frequency resonant converters accommodate great qualities such as improved safety with galvanic isolation, absence of high switching losses thanks to zero current (ZCS) and zero voltage switching (ZVS), along with better power density due to reduced volume of the magnetic components [10] [11].

Katarina Obradović is member of Department of Power Converters and Drives, University of Belgrade, 73 Bulevar kralja Aleksandra 73, 11020 Belgrade, Serbia (e-mail: obradovick15@gmail.com).

Emilija Lukić is member of Department of Power Converters and Drives, University of Belgrade, 73 Bulevar kralja Aleksandra 73, 11020 Belgrade, Serbia (e-mail: emilija.lukic505@gmail.com).

Jovana Plavšić is member of Department of Power Converters and Drives, University of Belgrade, 73 Bulevar kralja Aleksandra 73, 11020 Belgrade, Serbia (e-mail: plavsicjovana37@gmail.com).

Aleksandar Milić is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: milic.aleksandar@etf.bg.ac.rs).

Numerous resonant tank configurations are possible such as parallel, series or a combination of both resonant phenomenon [12]. The voltage gain of a series resonant configuration (SRC) can take a maximum of 1, while a parallel resonant converter (PRC) can achieve up to 10. Three and four reactive energy storage elements that combine parallel and series resonant phenomenon are more stable for a wide range of loads, and additionally parasitic inductances or capacitances featured in a converter can be employed [13]. Amongst them, LLC is widely used in today's DC/DC converters. Nevertheless, CLLC has its application in bidirectional power conversion systems, while LLC is efficient unidirectional resonant tank that has great efficiency under light load and therefore presents a better option than, for example, LCC for PV microinverter purposes [14].

Resonant converter with LLC tank configuration can be a MPPT stage of a two stage three phase microinverter [15]. The microinverter with galvanic isolation provides improved consumers' safety and can be appropriated for variety of PV systems' realizations [16] - [18]. Additionally, development of a dual input DC/DC converters with LLC resonant tank [19] can be considered because of its abillity to manage wide input voltage rate or it can form a scalable power conversion system as presented in [20].

This paper represents a brief overview of performance of the LLC resonant tank and design of resonant elements for a converter shown at Fig. 1 with 1000 W rated power. Full bridge is considered before resonant tank, while diode rectifier is on transformer's secondary side before DC link.

In the Section II voltage gain and load dependance of the LLC resonant tank is introduced. Following, design procedure of a tank is described with capacitor selection in Section III and inductor design in Section IV. In the last section, experimental results are presented.

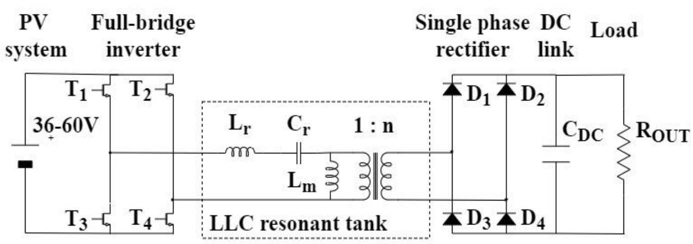


Fig. 1. Resonant LLC topology of a DC/DC converter with a passive load

II. LLC VOLTAGE GAIN AND LOAD DEPENDENCE

The design of a resonant tank requires input quantities' information including aforementioned rated power with 48 VDC input voltage as well as other parameters listed in Table I. While designing the presented system, the assumption was adopted that the output voltage of the converter will be maintained at a constant value by a proper control system.

Quality factor Q is of a great influence for the analysis of the tank. It represents the change in a converter's voltage gain curve at a different load conditions [21] and is characterized by equation (1)

$$Q = n^2 \cdot \frac{\pi^2}{8} \cdot \frac{P_{out}}{V_{out}^{2}} \cdot \sqrt{\frac{L_r}{C_r}} \qquad (1)$$

where P_{out} [W] is rated power, V_{out} [V] is an output voltage and n the ratio of secondary and primary number of turns. As it can be seen, Q factor is greater as the load approaches to the nominal value. Additionally, resonant frequency of the LLC resonant tank is described by

$$f_r = \frac{1}{2 \cdot \pi \cdot \sqrt{(L_r + L_m) \cdot C_r}} \qquad (2)$$

L_r being series resonant inductor, C_r series resonant capacitor and L_m the magnetizing inductance of the transformer.

Voltage gain of the tank is determined by the ratio of the switching and resonant frequency, load conditions (Q factor) and inductor ratio (m) as can be seen in the following equation

$$M = \frac{1}{\sqrt{\left[1 + \frac{1}{m} - \frac{1}{m} \cdot \left(\frac{f_s}{f_r}\right)^2\right]^2 + Q^2 \cdot \left(\frac{f_s}{f_r} - \frac{f_r}{f_s}\right)^2}}. \qquad (3)$$

Inductor ratio m is a constant that defines a ratio of a magnetizing inductance of a transformer L_m and series resonant inductance L_r. Increasing m value, voltage gain characteristics are more flattened thus more stable in wide frequency range, with reduced voltage peak on very light load. This effect can be seen on Fig. 2 where voltage gain characteristics are shown for a number of different Q factors and two different m values – high on the right and low one on the left.

In case m factor is high, LLC resonant tank condenses to a two-element circuit, meaning only L_r and C_r are utilized in a resonant occurrence and current through L_m is reduced. Alternating charging and discharging of the resonant capacitor and inductor leads to a reactive power exchange within these two reactive elements in the circuit which minimizes the apparent power of the tank. With reduced apparent power, the rms value of the current through the series resonant components is lowered and therefore the voltage drops, and Joule's power losses are lessened. High L_m is easily achievable and justified since transformer is there to provide galvanic isolation.

To make sure that larger deviation from the resonant frequency, due to switching frequency change, does not mitigate the voltage gain, the minimum Q factor must be established since that state could jeopardize elements bringing attenuation larger than one as seen in Fig. 2. This is because in case of light load, high values of the equivalent R_{out} could approach to similar value to $X_m = 2 \cdot \pi \cdot f \cdot L_m$ and parallel phenomena comes to the fore. The appropriate minimum Q value for the calculations must be determined knowing the expected load conditions. With all the requirements said, quality factor is taken to be 0.4.

TABLE I
INPUT PARAMETERS FOR THE DESIGN OF REACTIVE ELEMENTS

Parameter	Symbol	Value
Quality factor	Q	0.4
Turns ratio	n	11
Inductor ratio	m	100
Rated power [W]	P_{out}	1000
Input voltage range [V]	V_{in}	36 - 60
Output voltage [V]	V_{out}	450
Duty cycle	D	0.5

III. CAPACITOR SELECTION

In attempt to achieve high power density it is necessary to minimize volume of resonant elements by operating on high frequencies. However, it should be noted that working with

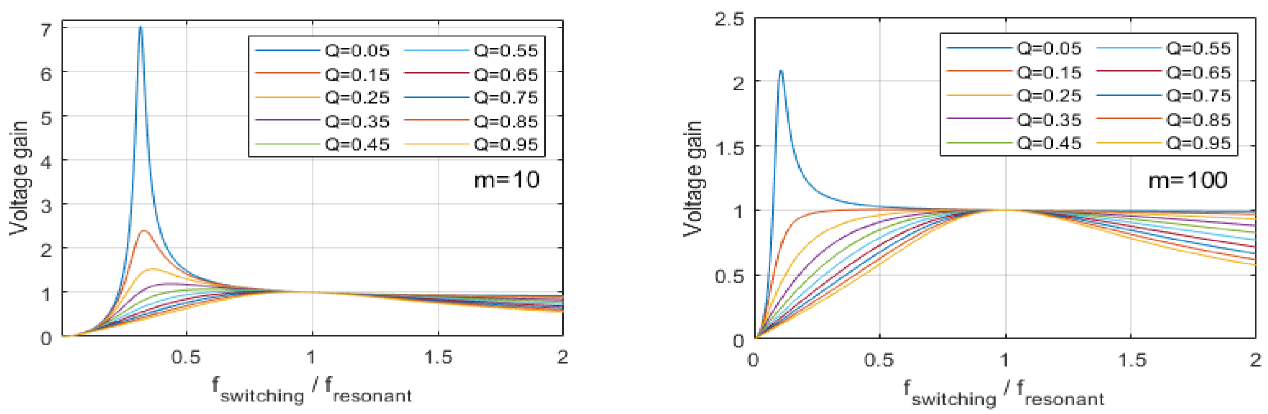


Fig. 2. LLC resonant tank voltage gain dependance on load conditions (Q) and switching frequency for two different values of m – 10 (left) and 100 (right)

too high frequencies brings other problems, for example higher switching losses on semiconductor devices. These losses are significantly lowered in LLC resonant topology due to zero voltage and zero current switching [22].

Values of resonant inductances and capacitance are related through resonant frequency f_r as show in equation (2). Capacitance as a function of frequency for agreed quality factor Q calculated using (1) and (2) is presented on Fig. 3.

An important step in the process of choosing resonant elements is to determinate frequency range in which the value of resonant capacitance stays unchanged. It has been noticed that after 150 kHz the curve portraited in Fig. 3 approximately stays in the flat region. After profound analysis of the transformer's power density, it is established that working with the frequency of 200 kHz with the chosen core material N87, volume of the transformer is sufficiently reduced.

Software simulations in LTspice were conducted applying switching frequency of 200 kHz, nominal input voltage 48 VDC, nominal active power 1 kW, capacitance value determined according to results based on Fig. 3, inductance value calculated by (2) for designated C and input parameters mentioned before. Based on simulation results in rated conditions, resonant current takes absolute value of 24 A.

From the manufacturer's documentation the one can find that an increase in resonant frequency leads to decreases in the maximum rated capacitor voltage as well as maximum rated current. Hence, further frequency rise would result in lowering capacitor's maximum rated current. It means usage of more parallel connected capacitors would be required in order not to pass maximum rated current of a single unit.

Based on the Fig. 3 for the frequency of 200 kHz, needed capacitance equals 1.327 µF. To maintain value of the capacitance for which the resonant current has the minimum value, LTspice simulations were performed. Capacitance value was changed in the range between 1 µF and 2 µF. It was necessary to secure that the current value in that range of capacitances does not fluctuate intensively for the input voltages. Minimum resonant absolute current value was 24 A, achieved in case of the capacitance value of 1.2 µF and inductance of 0.528 µH.

Since inductor will be designed according to the needs and considering the presented evaluation, selection criteria for capacitor are:

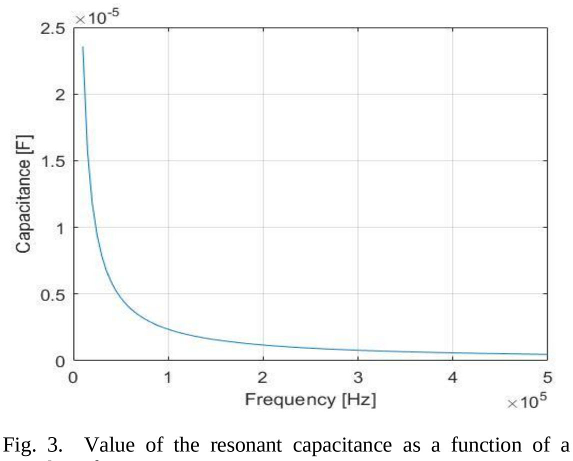


Fig. 3. Value of the resonant capacitance as a function of a resonant switching frequency

- absolute maximum current higher than 24 A on 200 kHz in case of using only one capacitor,
- low equivalent series resistance (ESR),
- capacitance value 1.2 µF,
- flat curve capacitance in a function of frequency,
- flat curve ESR in a function of frequency.

In case capacitor with these parameters is not available, possible solution would be to use more than one, i.e. x units with x times smaller capacitance rated for at least $1/x$ times of the mentioned current. However, suitable component was successfully found, so one unit can be used, and it is KEMET's film capacitor C4BSPBX4120ZB0J.

IV. Inductor design

After established inductance value of 0.528 µH, resonant inductor can be designed. Transformer's leakage inductance is not taken into consideration as a part of the equivalent inductance since it depends on various factors such as the tightness of the windings and windings' layout. Additionally, it is expected to have very low value.

Knowing the operating frequency, appropriate material can be selected according to manufacturer's suggestion. Inductor core's shape is taken to be E due to simplicity of winding and higher availability on the market. For 200 kHz suggested materials available for E cores are N87, N95 and N97.

A. Determining the core size

Firstly, according to [23], the core's dimensional parameter, S_wS_c, which represents product of the core's cross section (S_c) and the core's window area (S_w), should be above the value calculated using the following formula

$$S_wS_c = \sqrt[7]{\left(\frac{\sqrt{1+k_\gamma}\cdot K_i \cdot L \cdot I_{peak}^2}{B_{\max}\cdot K_t \cdot \sqrt{k_u \cdot \Delta T}}\right)^8} \cdot 10^{12} \qquad (4)$$

where:

- k_γ is the ratio between core and copper losses taken to be 1,
- K_i is the current waveform constant,
- L is the needed inductance value,
- I_{peak} is the maximum current intensity,
- B_{max} is the maximum value of flux density limited by the saturating flux density,
- K_t is the constant suggested by [23] to be 48200 as the core shape is not pot type; considering this constant acknowledges the copper resistivity on 20°C, it should be scaled so that the specific electrical resistivity is the one on 100°C which is $2.3\cdot10^{-8}$ Ωm,
- k_u is window utilization factor and
- ΔT is the core temperature rise.

Resulting S_wS_c values for each material are in Table II as well as the values of each listed parameter.

Inductance value for any core can be calculated as:

$$L = A_L \cdot N^2 \qquad (5)$$

where N is the number of turns. After selecting core that has S_wS_c value higher than the one from the Table II, A_L value can be read from the core's datasheet and L calculated using (5).

However, due to high current application, it is very important to take into consideration the fact that the diameter of the conductor will be relatively big. Therefore, prior to selecting the smallest possible core in order to achieve high power density, it is important to check if the window area is large enough for the turns.

B. Determining the size of the conductor

Current density should be below 4 A/mm^2 as a good practice, and as suggested by [24]. As current will not exceed 24 A, the diameter of the conductor should be at least 2.8 mm.

Owning to high frequency, it is advisable to use litz wires. This is due to small skin depth of the high frequency current which can be calculated based on the following equation [25]

$$\delta = \sqrt{\frac{\rho_{Cu}}{\pi \cdot f \cdot \mu}} \tag{6}$$

ρ_{Cu} being the specific electrical resistance of the copper, f operating frequency in Hz and μ material's permeability. Based on (6), for 200 kHz the strand of the conductor should be below 292 µm. Due to proximity effect, it appeared that smaller strands result in better efficiency, which has also been a suggestion by litz wires' manufacturer. Finally, based on the market available wires in required quantity, selected conductor has 2800 strands of the diameter 0.072 mm.

Taking the conductor's size into consideration besides the minimum S_wS_c value from the Table II, smallest available core that has large enough window area is the core E 25/13/7. It is produced in materials N87 and N97 and it has approximate A_L value 1850 nH and 1950 nH respectfully. Still, the inductance value computed using (5) even with only one turn is too high. This can be solved by adding a gap to the core.

TABLE II
INPUT PARAMETERS AND NEEDED S_wS_C VALUES

Symbol	Parameter	N97	N87	N95
B_{sat}	Saturating flux density [T]	0.41	0.39	0.41
L	Inductance [µH]	0.52		
K_i	RMS and peak current ratio	0.707		
K_t	Core shape constant	41703		
I_{peak}	Current peak value [A]	42.43		
k_u	Window utilization factor	0.8		
ΔT	Temperature rise [°C]	70		
B_{max}	Maximum flux density [T]	0.4	0.38	0.4
S_wS_c	Core size [mm^4]	529.3	561.2	529.3

C. Determining the gap size

Inductance of the gapped core can be calculated as

$$L = \frac{N^2}{\frac{1}{\mu_0} \cdot \left(\frac{l_g}{S_c} + \frac{1}{\mu_r} \cdot \frac{l_c - l_g}{S_c} \right)} = \frac{N^2}{\frac{1}{\mu_0 \cdot \mu_{eff}} \cdot \frac{l_c}{S_c}} \tag{7}$$

where l_g [mm] is the gap length, l_c [mm] the length of the flux path corresponding to the chosen core, μ_r is the material's relative permeability, μ_0 [H/mm] magnetic permeability of the air and S_c [mm^2] core's cross section. Thus, adding the air gap μ_{eff} is being shrunk and therefore N has achievable value. Additionally, air gap contributes to stability of the magnetic properties of the core [26].

The core E 25/13/7 in material N97 cannot be gapped by the manufacturer, so the chosen material is N87. The manufacturer has predefined values for air gaps: 0.1 mm, 0.16 mm, 0.25 mm, 0.5 mm, and 1 mm. In Table III it is listed how many turns is needed in case of different gap lengths. Also, the A_L value in case fringing effect is taken into consideration is given by the manufacturer and presented in Table III.

For the 0.5 mm air gap the difference in desirable and resulting A_L and μ_{eff} is the least as seen from Table III. Hence, the chosen core is E 25/13/7 in material N87 with the air gap of 0.5 mm having two turns of litz wire consisting of 2800 strands of the diameter 0.072 mm. Resulting inductance is expected to be 0.5976 µH.

TABLE III
GAPPED E 25/13/7 CORE ANALYSIS

Air gap length [mm]		0.1	0.5	1
Number of turns		1	2	3
Desired μ_{eff} value		462	115	51
Desired A_L value [nH]		530	132.5	58.9
μ_{eff}	No fringing effect	456	109	56
	With fringing effect	422	130	78
A_L [nH]	No fringing effect	523.2	125.4	64.3
	With fringing effect	484.4	149.4	90.0

D. Finite Elements Method Analysis

Using software tool FEMM 4.2 finite elements method analysis is conducted. The graphical result is shown in Fig. 4. It should be noted that the flux density in the fringes passes the saturating point.

Inductance value calculated by the software is 0.086 µH. However, this value should be scaled since FEMM 4.2, while calculating equivalent inductance, considers only the part of the turn that is inside the core window. Thus, the value given by the program should be multiplied with the ratio of the turn length in millimeters and doubled core depth, i.e. 14 mm. Corrected value equals to 0.307 µH.

V. EXPERIMENTAL RESULTS

The main intent of testing is to observe the behavior of the resonant topology. Prototype of LLC resonant DC/DC

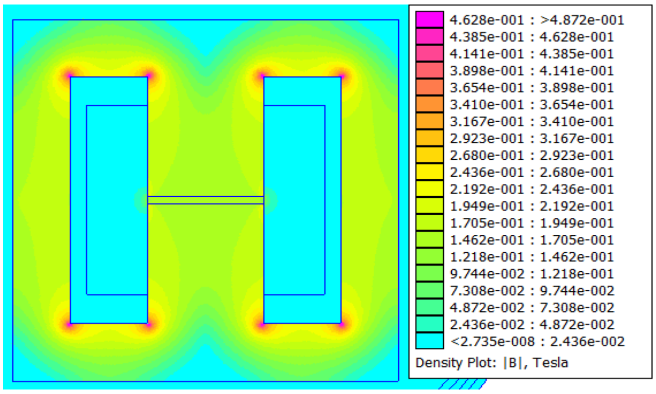


Fig. 4. Flux density in the core done in the software tool FEMM 4.2

converter for 1 kW was developed in Digital Drive Control Laboratory in School of Electrical Engineering, University of Belgrade.

Designed prototype with testing circuit is presented in Fig. 5. Prototype was realized with gallium nitrite transistors as input DC to AC converter in a full bridge configuration using two EPC1022 development boards. However, since these transistors could work with even higher frequencies and remain satisfactorily efficient, it is advisable to, but not limited to, consider them in 500 kHz or higher application. Nevertheless, they can function properly on 200 kHz as well, so they had been taken as a part of the prototype, since accent is on the resonant tank rather than full-bridge parameters.

The measured parameters of the designed and selected resonant tank elements are 0.648 μH for resonant inductance, 49.55 μH for transformer's magnetization inductance, and capacitances of 100 μF and 1.2 μF for DC link and resonant capacitor respectively with load resistance of 650 Ω. Transformer's magnetization inductance, series inductance and capacitance were measured using LRC meter, with 200 kHz excitation signals. For driving transistors in full-bridge configuration, Texas Instruments launch pad 28379D was used employing Power Width Modulation (PWM) signals with adjustable dead time duration. Soft start was necessary to implement in order to avoid unwanted impulsive changes of resonant current. It was realized by phase shift modulation of driving signals. During 3.6 s, one of PWM signal's duty cycle was changed in the range of 0 – 50 %.

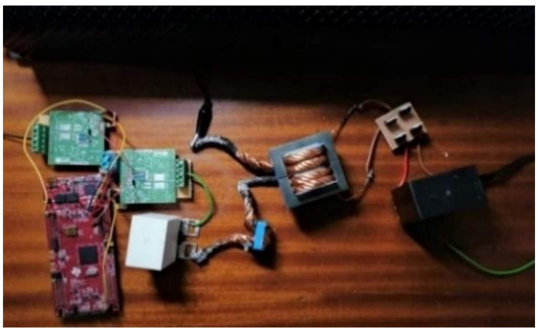


Fig. 5. Experimental setup of LLC resonant DC/DC converter

According to theoretical calculations for resonant tank parameters, switching frequency was set to 200 kHz. The voltage waveforms on primary and secondary transformer side with switching frequency of 200 kHz and 15 VDC at the input converter terminals are shown in Fig. 6. It can be noticed that the transformer's low side voltage is multiplied by the turns' ratio. Also, voltage gain of resonant tank is equal to one, as expected.

Waveform of resonant current is presented in Fig. 7 where input parameters were the same as mentioned before. Based on measurements, resonant inductance takes value of 0.648 μH instead of theoretical 0.528 μH due to imperfections of the inductor's development process. Resonant frequency was now estimated using (2) once again and set to 187 kHz.

Capacitances on the output of the diodes of the rectifier along with output capacitances of the transistors from full-bridge are not being fully discharged during the dead time. As a result, these parasitic capacitances and the circuit inductance had started to resonate, and flinches appeared. In Fig 8. these oscillations on the resonant current waveform can be noticed.

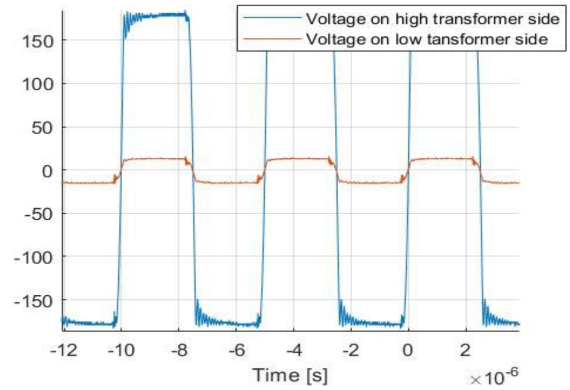


Fig. 6. Voltages on low and high transformer side with 15VDC input voltage and switching frequency of 200kHz.

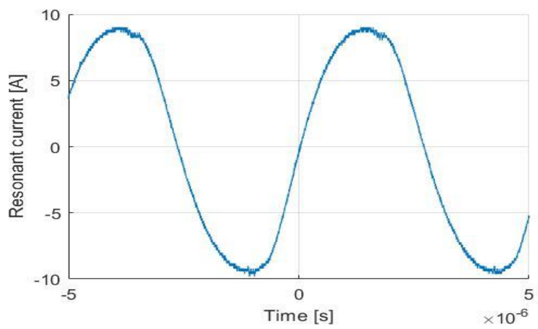


Fig. 7. Waveform of resonant current for 15 VDC input voltage and 200 kHz switching frequency

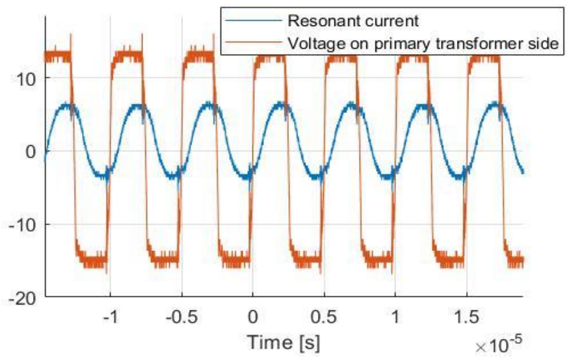


Fig. 8. Waveforms of resonant current and voltage on primary transformer side for 187 kHz as a switching frequency

VI. Conclusion

Having the advantage of reduced switching losses due to zero current and zero voltage switching, resonant DC/DC topologies are worth considering in the high frequency high current application in PV inverters. Numerous resonant tank topologies are possible, where LLC is the most common one in the terms of voltage gain stability and design simplicity.
In this paper the process of determining needed series capacitance and inductance's values from LLC resonant tank is presented. Transformer's magnetization inductance has 100 times larger value, so its impact has been neglected.

Thereafter, the procedure of choosing capacitor and inductor design is elaborated in detail.

On the developed prototype, waveforms of relevant variables were observed in a resonant circuit in order to verify the presented design procedure. Experimental tests have shown that the presented method gives satisfactory results in terms of the system's behavior. The resonant LLC converter, in addition to the possibility of working in MPPT mode and galvanic isolation, provides high efficiency and power density, thus reducing the volume and increasing the total efficiency of the conversion system. Due to aforesaid advantages, the presented resonant tank topology is a good choice for the needs of microinverter PV applications.

Further research could take into consideration other core shapes for inductor, include transformer's leakage inductance into calculations, as well as the analysis of higher working frequencies suitable for wide-bandgap devices.

ACKNOWLEDGMENT

We would like to express special gratitude to Efficient Power Conversion (EPC) company. Their generous donation of EPC1022 development boards has helped us to undertake this project and present experimental research that we have accomplished in the Laboratory.

References

[1] R. W. Erickson, D. Maksimović in "Principles of Steady State Converter Analysis" *Fundamentals of Power Electronics*, New York, United States of America: Soringer, 2001, ch. 2, sec 3. pp. 22-27

[2] R. W. Erickson, D. Maksimović in "Steady-State Equivalent Circuit Modeling, Losses, and Efficiency" *Fundamentals of Power Electronics*, New York, United States of America: Soringer, 2001, ch. 3, sec 4. pp. 50-52

[3] S. N. Soheli, G. Sarowar, M. A. Hoque and M. S. Hasan, "Design and Analysis of a DC -DC Buck Boost Converter to Achieve High Efficiency and Low Voltage Gain by using Buck Boost Topology into Buck Topology," *2018 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE)*, 2018, pp. 1-4

[4] S. Cuk, R. D. Middlebrook, "A new optimum topology switching DC-to-DC converter" in *IEEE Power Electronics Specialist Conference*, Palo Altro, CA., USA, 1977.

[5] Y. Konishi, Y.-F. Huang, M.-J. Hseih, "Utility-interactive high-frequency flyback transformer link three-phase inverter fir photovoltaic AC module" in *35th Annual Conference of IEEE Industrial Electronics*, Porto, Portugal, 2009.

[6] Z. Chen, Q. Wu and Y. Yuan, "A Novel Zero-Voltage-Switching Push–Pull High-Frequency-Link Single-Phase Inverter," in *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 4, no. 2, pp. 421-434, June 2016.

[7] A. Hillers, D. Christen and J. Biela, "Design of a Highly efficient bidirectional isolated LLC resonant converter," 2012 15th International Power Electronics and Motion Control Conference (EPE/PEMC*)*, 2012.

[8] K. Siebke and R. Mallwitz, "Operation Mode Analysis of the CLLC Resonant Converter," *2019 IEEE 13th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG)*, 2019, pp. 1-6

[9] Y. Zhuang, F. Liu, X. Zyang, X. Diao, J. Jiang and J. Sun, "Direct Frequency Control Based MPPT Algorithm of LLC Resonant Converter for Photovoltaic System" in *2019 IEEE Energy Conversion Congress and Exposition (ECCE)*, Baltimore, Maryland, USA, 2019.

[10] Hillers, A., Christen, D., & Biela, J. (2012). *Design of a Highly efficient bidirectional isolated LLC resonant converter. 2012 15th International Power Electronics and Motion Control Conference (EPE/PEMC)*, 2012.

[11] F. Musavi, M. Craciun, M. Edington, W. Eberle, and W. G. Dunford, "Practical design considerations for a LLC multi-resonant DC-DC converter in battery charging applications," in Proc. 27th Annu IEEE APEC Expo., 2012, pp. 2596–2602

[12] R. L. Steigerwald, "A comparison of half-bridge resonant converter topologies," in *IEEE Transactions on Power Electronics*, vol. 3, no. 2, pp. 174-182, April 1988.

[13] R. Severns, "Topologies for three element resonant converters," Fifth Annual Proceedings on Applied Power Electronics Conference and Exposition, 1990, pp. 712-722

[14] S. Mao, J. Popovic, R. Ramabhadran and J. A. Ferreira, "Comparative study of half-bridge LCC and LLC resonant DC-DC converters for ultra-wide output power range applications" *2015 17th European Conference on Power Electronics and Applications (EPE'15 ECCE-Europe)*, 2015, pp. 1-10

[15] Alamir, O. Abdel-Rahim, M. Orabi and M. Ismeil, "Two-Stage Resonant Three-Phase Micro-inverter for Grid-Tie PV Application," *2019 IEEE Conference on Power Electronics and Renewable Energy (CPERE)*, 2019, pp. 459-464

[16] G. Rubino, L. Rubino, N. Serbia, P. Ladoux and P. Marino, "LLC resonant converters in PV applications comparison of topologies considering the transformer design," *2013 International Conference on Clean Electrical Power (ICCEP)*, 2013, pp. 37-41

[17] C. Buccella, C. Cecati, H. Latafat and K. Razi, "A grid-connected PV system with LLC resonant DC-DC converter," *2013 International Conference on Clean Electrical Power (ICCEP)*, 2013, pp. 777-782

[18] A. Bousselham and A. Elrayyah, "Control and efficient operation of low cost multi PV-based LLC resonant microinverters," *2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC)*, 2016, pp. 3457-3462

[19] S. M. Tayebi, H. Hu, O. Abdel-Rahman and I. Batarseh, "Design and analysis of a dual-input single-resonant tank LLC converter for PV applications," *2018 IEEE Applied Power Electronics Conference and Exposition (APEC)*, 2018, pp. 476-483

[20] C. Schaef and J. T. Stauth, "Multilevel Power Point Tracking for Partial Power Processing Photovoltaic Converters," in *IEEE Journal of Emerging and Selected Topics in Power Electronics*, Dec. 2014, vol. 2, no. 4, pp. 859-869

[21] G. Hsieh, C. Tsai and S. Hsieh, "Design Considerations for LLC Series-Resonant Converter in Two-Resonant Regions," *2007 IEEE Power Electronics Specialists Conference*, 2007, pp. 731-736

[22] X. Zhou, Z. Liang and A. Huang, "A High-Dynamic Range Current Source Gate Driver for Switching-Loss Reduction of High-Side Switch in Buck Converter," in IEEE Transactions on Power Electronics, June 2010. vol. 25, no. 6, pp. 1439-1443

[23] W. G. Hurley, W. H. Wölfle, "Inductor Design," in *Transformers and Inductors for Power Electronics*, Chichester, United Kingdom: John Wiley & Sons Ltd., 2013, ch. *3*, sec. *1-2*, pp. *61-64*

[24] A. I. Pressman, K. Billings, T. Morey, "Transformers and Magnetics Design," in Switching Power Supply (3th Edtion), New York, United States of America: McGraw Hill Education, 2013, ch. *7*, sec. 3, p. *313*.

[25] A. V. Bossche, V. C. Valchev, "Soft Magnetic Materials" in *Inductors and Transformers for Power Electronics*, Boca Raton, United States of America: CRC Press Taylor & Francis Group, 2005, ch. *3*, sec. *3*, pp. *136–137*

[26] W. G. Hurley, W. H. Wölfle, "Inductance," in *Transformers and Inductors for Power Electronics*, Chichester, United Kingdom: John Wiley & Sons Ltd., 2013, ch. *2*, sec. *1*, pp. *25–38*

Implementation and testing of basic algorithms in PV systems with batteries on a common DC link

Katarina Ćeranić, Mila Gligorijević, Lazar Stojanović and Aleksandar Milić

Abstract—**Photo-voltaic systems with batteries on a common DC link, i.e. the concept of the Point of Common Coupling (PCC), is increasingly in use. In such systems, it is necessary to achieve the basic system functionalities, such as bidirectional battery operation, efficient Maximum Power Point Tracking (MPPT) regimes on the PV panels, as well as constant voltage on the common DC link for the consumer needs. In this paper, the basic algorithms for battery and panel operation in the MPPT mode are provided. The analysis was first verified in software packages and later by implementing algorithms on the developed low-power prototype of the system, where the basic functionalities were presented. Additionally, the robustness of the algorithms to power transients and disturbances which are common in the PCC systems was tested.**

Index Terms—**DC link, Point of Common Coupling, battery, MPPT, battery charging strategy.**

I. Introduction

One of the leading causes of global warming is production of the electrical energy. Commonly used energy sources are non-renewable, such as coal, oil and natural gas, where their usage directly results in the pollution of air, water and land. In the past few decades, the use of renewable energy sources, especially wind and solar power, has become more common [1][2]. Further, a stand-alone PV system whose application context is specific to the countryside or isolated locations for self-feeding, is seen as a substitute for the utility grid connection [3]. In such applications, photo-voltaic systems (PV) almost always imply the use of batteries in their continuous operation. Namely, this application enables the consumer to be independent from the electrical grid, as well as, to simultaneously act as a consumer and a producer of electrical energy. Thus, there is a need for a scaled down energy storage system which interacts with the clean energy source.

Katarina Ćeranić is a student at Power Converters and Drives Department, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: ceranic.katarina@gmail.com).

Mila Gligorijević is a student at Power Converters and Drives Department, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: mila.gligorijevic7@gmail.com).

Lazar Stojanović is a student at Signals and Systems Department, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: lazastojanovic99@gmail.com).

Aleksandar Milić is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: milic.aleksandar@etf.bg.ac.rs).

The main disadvantage of PV panel employment is the fluctuating output power induced by the variable solar irradiance. This introduces difficulty into fully capitalizing on the panel production, hence creating a need for a control algorithm which would track the maximum power point – MPPT algorithm [4], such as Perturb and Observe (P&O), Incremental Conductance (IC), Current Sweep (CS) technique [5][6], or Open Circuit Voltage (OCV) technique [7][8], along with the Particle Swarm Optimisation (PSO) based algorithms [9]. The main goal of these algorithms is circumventing the practical issues which arise with the PV panel implementation - variable irradiance or damaged cells. Comparison of these algorithms has shown that each has its advantages and disadvantages – where P&O and IC implementation is straightforward, their robustness to disturbance is not a strong point. These two algorithms are also incapable of locating the global maximum on the P-V curve, which usually occurs when the PV string has multiple bypass diodes i.e. multiple power peaks. PSO, in particular, has found its implementation in large string operations, where multiple bypass diodes are unavoidable. However, more advanced algorithms, PSO and CS, are much more complex with their benefits being insignificantly greater than those of P&O and IC in terms of microcontroller implementation. The main shortcoming of the VOC technique is the fact that it requires the power delivery to be halted every time the open circuit voltage of the PV string is needed for further calculation.

Although the development of MPPT algorithms as well as controlled battery charging and discharging algorithms have progressed, there are applications where basic algorithms, reliably implemented, can achieve acceptable results. There, fast high-performance processors are redundant and, thus, avoided, which immediately reduces the overall cost of the product. For this reason, this paper deals with the implementation and testing of basic algorithms in the system presented in Fig. 1, which contains PV panels and a battery connected to a common point. The paper is organized as follows – in Section II an overview of the analyzed system is presented, section III provides basic analysis of the implemented control algorithms. Further, the results of the comparison of the applied algorithms obtained in the simulations and in the experimental setup are presented in Sections IV and V, respectively. Finally, a brief conclusion of the comparison is given.

II. System overview

The system, on which the control has been implemented, falls in line with the streamlined topology – it is composed of a buck converter as the input stage and a four-switch synchronous buck-boost converter connected to the battery [10]. The point of common coupling (PCC) for these two converters is the DC link, which is also the connection point for the consumer. The prototype parameters are provided in Tables I and II, and the entire power stage schematic is shown in Fig. 1.

TABLE I
System parameters

	L[μH]	C[μF]	R_{DSon}[Ω]
Buck	11.3	22	0.01
Buck-Boost	4.7	100	0.008
DC Link	/	450	/

TABLE II
PV panel parameters

V_{OC}[V]	I_{SC}[A]	V_{MPP}[V]	I_{MPP}[A]
21.9	1.84	17.6	1.7

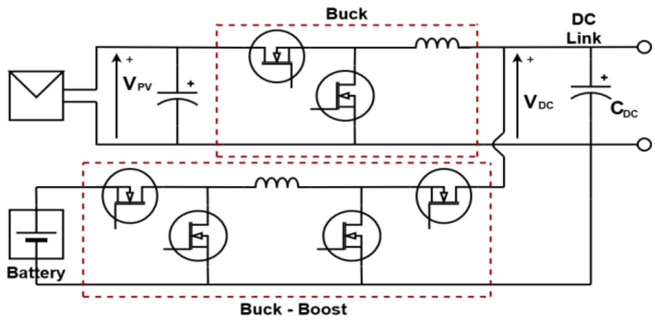


Fig 1. System overview

The hardware design which is increasingly represented in the industry comprises of an input stage communicating with the PV panels and a battery power stage where the common point is the DC link. The DC link voltage is kept consistently within the desired margins, either by generating the power from the PV panels or by discharging the battery [11]. Li-ion battery cells have become the mainstream portable power source due to their exceptional characteristics, such as high specific power density, high cell voltage, low self-discharging rate, high charging current and long life cycle [12]. Hereof, the developed prototype, which will be discussed in more detail in Section V, uses two identical Li-ion batteries connected in series and PV panels whose characteristics are shown in Fig. 2.

In general applications pertaining to higher power derivation, the input step would have been a boost converter, due to its continuous input current [13]. This paper is considering low-voltage low-power applications, so a buck converter as the input stage was chosen, despite the discontinuous input current, which ought to render the MPPT mode inoperable. The desired MPPT operation regime is feasible if a capacitor is connected in parallel to the panels, as it enables the continuous power necessary for the MPPT algorithms. Thus, the input buck converter lowers the panel voltage to the DC link voltage and the MPPT algorithm governs the duty cycles of the input buck. The battery buck-boost keeps the DC link voltage constant and charges or discharges the battery in accordance with the system state [10].

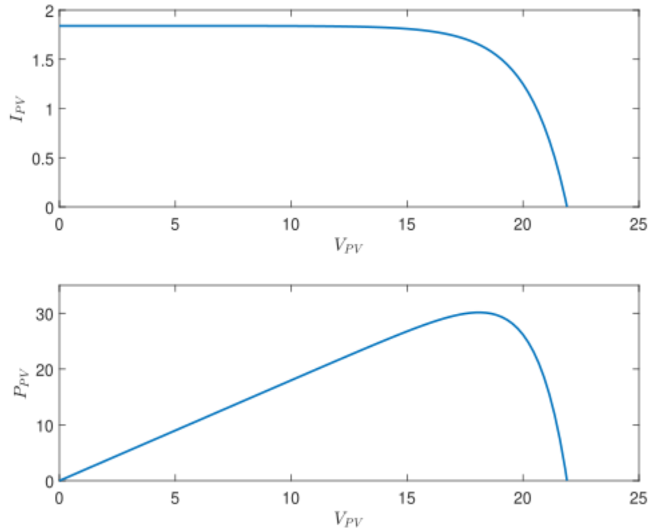


Fig. 2. Current to Voltage and Power to Voltage Characteristics

Various converters which allow the battery management i.e. bidirectional power flow have been devised. Some of them are: buck-boost, SEPIC, Cuk, synchronous four-switch buck-boost and multi-level bidirectional converters [14]. The chosen converter here is the synchronous four-switch buck-boost due to its ease of implementation.

III. Analysis of control algorithms

A. MPPT algorithms analysis

In this Section, the two employed MPPT algorithms are the "hill climbing" algorithms P&O and IC – shown in Fig. 4 and Fig. 5. P-V characteristics of the panels, which are used in both algorithm designs, are shown in Fig. 2. These two algorithms have been chosen for their ease of implementation and low computational power they expend. P&O and IC both slide the operating point along the P-V characteristic of the panels so that the string produces the maximum power available at any given point in time [15]. When using P&O, the controller incrementally adjusts the harvested power by measuring the panel voltage and panel current, and then either takes a step forward or backward along the P-V curve depending on the needs of the system, as shown in Fig. 3. The main issue with this method is that even when the input power is stable, the derived power oscillates in the vicinity of the MPP.

IC does not produce output power oscillations. It relies on the slope of the P-V curve and, as such, upon reaching the MPP, remains there due to zero inclination. The controller maintains these working states until the irradiation changes.

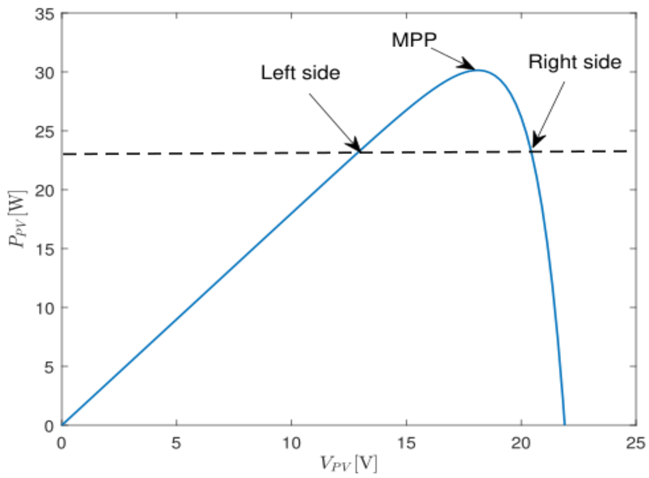


Fig 3. P-V operating point positions

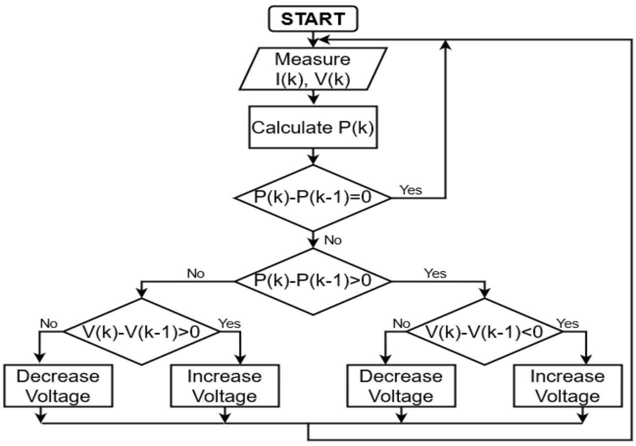


Fig 4. Perturb and Observe Algorithm

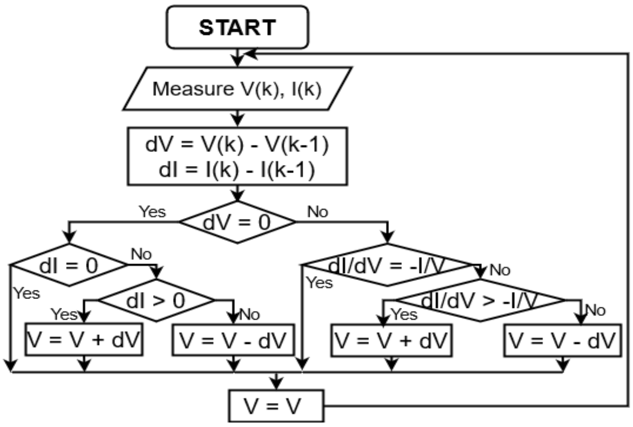


Fig 5. Incremental Conductance Algorithm

B. Battery charging

Another topic of importance is the battery charging strategy. In order to charge, the battery needs to be connected to a power supply. This triggers an oxido-reduction reaction – oxidation occurs on the positive electrode, the one releasing electrons, while reduction manifests on the negative, attracting the released electrons and charging the battery. Discharge happens when the battery is connected to a load – the process of discharging is directly opposite to charging [16]. The battery charging process is done by a combination of Constant Current (CC), typically no greater than 1C or 2C, and Constant Voltage (CV) operating modes, as show in Fig. 6 [12].

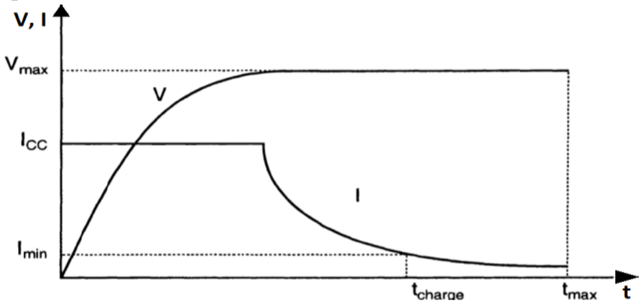


Fig 6. CC-CV Battery Charging

In practice, the values of voltage and charging current are dictated so as to preserve the cell lifespans – high values of maximum voltage and their application for extended periods of time should be avoided [17][18]. Since the prototype operates with the above-mentioned batteries, which are Li-ion based, the best charging method is Constant Current – Constant Voltage (CC-CV). Constant Current (CC) entails charging with a constant current value. This is maintained until the battery reaches the designated voltage value. So as to not overcharge the battery, the voltage value at which CC is stopped has to be less than 100% - usually the cut-off voltage is around 80-90%. In the developed prototype, the charge rate is limited to 1C. This process is then followed by Constant Voltage (CV) charging mode – it charges the battery to full capacity while the current exponentially reduces.

C. CC-CV and the stability of operating points

The CC-CV method charges a battery with a constant current until the battery voltage increases to the constant voltage limit. Then the voltage is kept at said limit while the charging current gradually decreases. In the CC-CV charging method, the CV regime typically lasts longer, prolonging the total charging time [19]. By employing the optimal constant values of current and voltage for the used battery, one is able to achieve the most efficient battery charging process.

Another important operational mode pertains to the situation where the load power consumption is low and the battery is already full. Therefore, it is necessary to move from the MPP in order to not overcharge the battery or jeopardize the DC link voltage. Assuming that the eventual load and the battery combined are able to consume less power than what the panel can generate, in these cases, it is necessary to move the operating point to the part of the curve where less power is generated. This is done by controlling the input power via the input buck converter and is the main idea behind the CC-CV algorithm implemented in this paper.

Bearing in mind the topology of the panel converter, the peak in power production is going to firstly be reflected in the DC link voltage V_{DC} – which is going to start increasing. Each power demand corresponds to two operating points – one to the left of the MPP, and another one to the right, refer to Fig. 3. The criteria for the operating mode choice is the system stability, which is ascertained through a system model. Assuming that the converter is in equilibrium, the operating point is tested to small perturbation. In (1) and (2) U_{CDC} represents the DC link capacitor energy. P_{PV} is the panel power and P_0 is the sum of all losses, load and battery consumption. The equation (3) is valid due to the nature of the converter – here, D refers to the duty cycle.

$$\frac{dU_{C_{DC}}}{dt} = P_{PV} - P_0 \tag{1}$$

$$U_{C_{DC}} = \frac{C_{DC} \cdot V_{DC}^{\;2}}{2} \tag{2}$$

$$V_{PV} = \frac{V_{DC}}{D} \tag{3}$$

Combining equations (3) and (2) and inserting that into (1) results in the following model:

$$\dot{V}_{PV} = \frac{1}{C_{DC} \cdot D^2 \cdot V_{PV}} \cdot (P_{PV} - P_0) \tag{4}$$

$$V_{PV} = f(V_{PV}) \tag{5}$$

To apply indirect Lyapunov method [20], the model (4) has to be linearized first:

$$\frac{\partial \dot{V}_{PV}}{\partial V_{PV}} = \frac{1}{C_{DC} \cdot D^2} \cdot \left(\frac{\partial P_{PV}}{\partial V_{PV}} \cdot \frac{1}{V_{PV}} - \frac{P_{PV} - P_0}{V_{PV}^{\;2}} \right) \tag{6}$$

In order for the equilibrium point $P_{PV}=P_0$ to be stable:

$$\frac{1}{C_{DC} \cdot D^2} \cdot \frac{\partial P_{PV}}{\partial V_{PV}} \cdot \frac{1}{V_{PV}} < 0. \tag{7}$$

Since $C_{DC}>0$, $V_{PV}>0$, (7) is equal to:

$$\frac{\partial P_{PV}}{\partial V_{PV}} < 0 \tag{8}$$

This is true only for the part of the P-V curve that is on the right side of the MPP, as seen in Fig. 3. A detailed overview of the employed CC-CV algorithm is provided in Fig. 7.

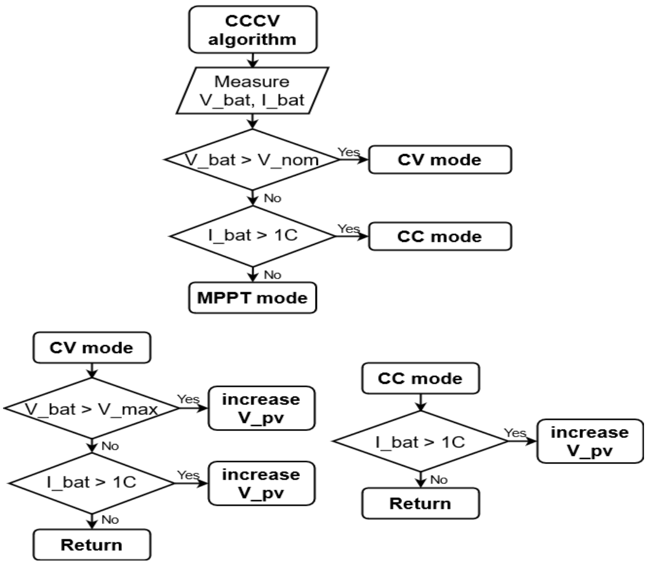


Fig. 7. CC-CV algorithm

IV. Simulation results

For the purpose of this paper, the analyzed system presented in Fig. 1 was simulated in a MATLAB - Simulink software package. In the simulations, it was adopted that the maximum output power of the system is 80 W, which corresponds to the developed prototype of the system. This is going to be discussed in more detail in the next chapter. The simulation conditions include irradiance which changes throughout the duration of the simulation and a load that can be optionally connected or disconnected. The MPPT algorithm adapting to variable irradiance is shown in Fig. 8. During the test, it was adopted that there is a consumer of constant power on the common DC link, which is at constant voltage. For this reason, from the presented results, one can identify a change in battery current depending on the power delivered from the panels. That is, depending on the generated power, the battery switches from the discharge mode into charging mode and vice versa. Also, it can be concluded that the implemented algorithm has a fast response with no switching upon varying the input irradiance.

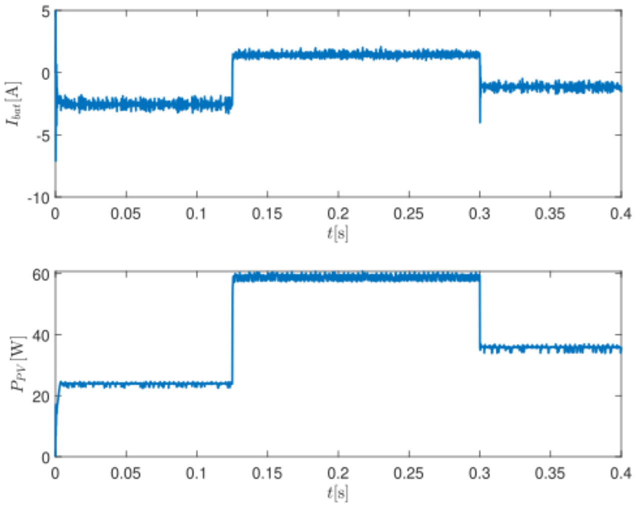


Fig. 8. MPPT mode with variable irradiance

Further, Fig. 9 depicts the battery charging current and the power derived from the panels when the load is suddenly disconnected during the CC charging mode. The current peaks, at the moment of disconnection, however, the controller quickly limits it to the maximum allowed value. The derived panel power is appropriately lowered, showcasing the controller's ability to adapt to disturbances. However, oscillations can be observed in the generated power of the panel, which in this mode, is delivered directly to the battery.

Similarly, the controller adjusts to the load disconnection during CV mode - Fig 10. The battery voltage is swiftly returned to the designated voltage value and the power is correspondingly reduced which is achieved by changing the active pulse width when controlling the switches on the panel converter. Similar oscillations can be observed in the output power of the panel as in the previous test.

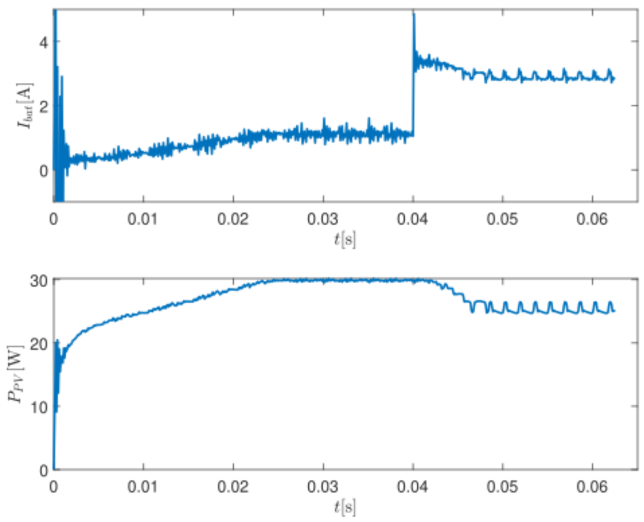


Fig. 9. CC mode with sudden load disconnection

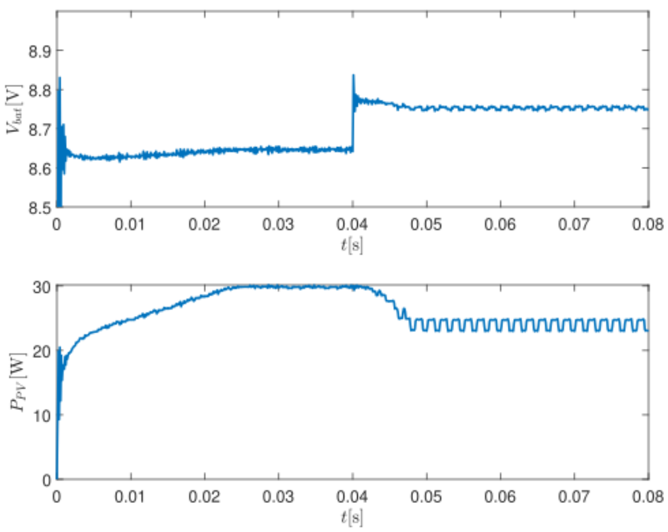


Fig. 10. CV mode with sudden load disconnection

V. EXPERIMENTAL RESULTS

For the required testing of the presented algorithms in actual working conditions, a prototype of the analyzed system was developed in the Digital Drive Laboratory, University of Belgrade. The experimental setup and the prototype can be seen in Fig. 11 and Fig. 12, respectively. The setup consists of two PV panels mounted on the wall. Four halogen lights supplied through two autotransformers simulate variable irradiation. The setup also includes the prototype, an oscilloscope and current probe monitoring the battery current. A resistive load is connected to the DC link. The presented algorithms were implemented on a low-power STM32l4r5qi processor with a code execution time of 10 µs.

Regarding the MPPT stage, tests of the two presented algorithms were performed. The results are presented in Fig. 13, where the MPPT rise time can be observed to be around 180 ms, as well as the power oscillation when the MPP is located. Theoretically, IC is supposed to be more stable once in the MPP, however, in this application, IC operates less effectively due to the induced noise on the prototype.

Fig. 11. Testing experimental setup

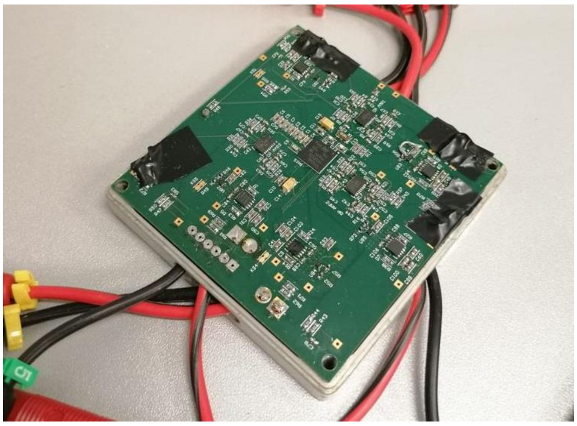


Fig. 12. Developed prototype

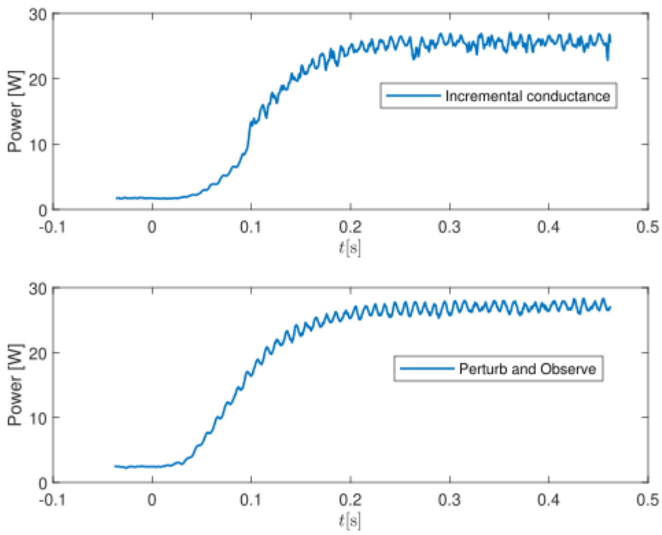


Fig. 13. MPPT methods comparison

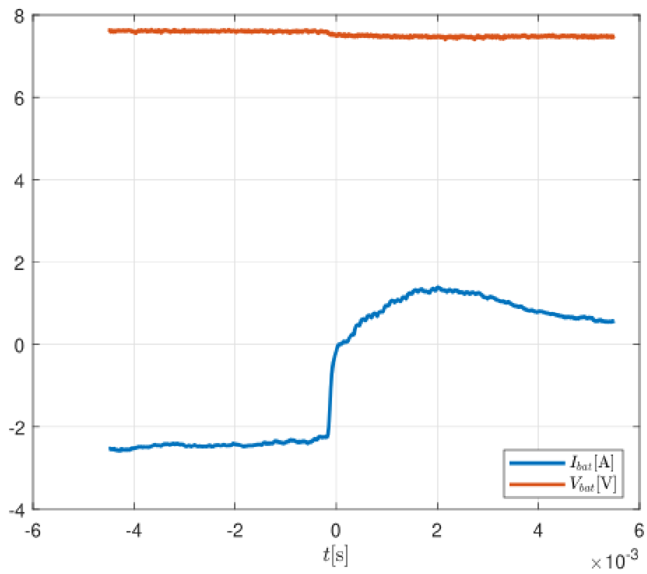


Fig. 14. Transition from charge to discharge

As for the battery, Fig. 14 shows the transition between charging and discharging modes as well as the DC link voltage. From the presented results, it can be concluded that the DC link voltage does not experience too many disturbances during this transient. This is achieved by an external voltage control loop of the converter on a battery that maintains the DC link voltage at a constant value of 7.8 V. An internal current control loop maintains the battery charging and discharging current. The switching frequency of 300 kHz has resulted in the low values of the charge and discharge current ripple. Also, it can be seen that both battery modes are successfully achieved – charging when the current takes a positive value and discharging when the current becomes negative. Insufficient input power is simulated by lowering the irradiance, which results in triggering the change from charge to discharge.

VI. Conclusion

This paper presents a simple design of an algorithm for controlling a low-power PV system supported by batteries. The logistics of the algorithms that are often encountered in practice are presented as they have been implemented on the developed prototype of the system. The simulation results were expanded upon by the experimental results. Furthermore, the three basic system requirements were met, bidirectional battery operation, panel operation in the MPPT mode as well as constant DC link voltage for the needs of powering an arbitrary consumer. The presented algorithms are quickly and easily implemented on low-power and low-cost processors. Of course, the presented solution has an extensive space for improvement when considering system performance, algorithm execution speed as well as robustness to disturbances.

References

[1] S. Krauter, "Greenhouse gas reduction by PV," 3rd World Conference onPhotovoltaic Energy Conversion, 2003. Proceedings of, 2003, pp. 2610-2613 Vol.3

[2] Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [B. Metz, O.R. Davidson, P.R. Bosch, R. Dave, L.A. Meyer (eds)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA

[3] F. Locment, M. Sechilariu and I. Houssamo, "Batteries and DC charge control of stand-alone photovoltaic system. Experimental validation," Proceedings of 14th International Power Electronics and Motion Control Conference EPE-PEMC 2010, 2010, pp. T12-43-T12-48

[4] Patil S.N. and R. C. Prasad, "Design and development of MPPT algorithm for high efficient DC-DC converter for solar energy system connected to grid," 2015 International Conference on Energy Systems and Applications, 2015, pp. 228-233

[5] T. Esram and P. L. Chapman, "Comparison of Photovoltaic Array Maximum Power Point Tracking Techniques," in IEEE Transactions on Energy Conversion, vol. 22, no. 2, pp. 439-449, June 2007

[6] Bodur, Mehmet; Ermis, M. (1994). "Maximum power point tracking for low power photovoltaic solar panels". Proceedings of the 7th Mediterranean Electrotechnical Conference: 758–761

[7] Ferdous, S.M.; Mohammad, Mahir Asif; Nasrullah, Farhan; Saleque, Ahmed Mortuza; Muttalib, A.Z.M.Shahriar (2012). 2012 7th International Conference on Electrical and Computer Engineering. ieee.org. pp. 908–911

[8] Hlaili, M., \& Mechergui, H. (2016). Comparison of Different MPPT Algorithms with a Proposed One Using a Power Estimator for Grid Connected PV Systems. International Journal of Photoenergy, 2016, 1–10

[9] Y. Liu, D. Xia and Z. He, "MPPT of a PV system based on the particle swarm optimization," 2011 4th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT), 2011, pp. 1094-1096

[10] M. Orellana, S. Petibon, B. Estibals and C. Alonso, "Four Switch Buck-Boost Converter for Photovoltaic DC-DC power applications," IECON 2010 - 36th Annual Conference on IEEE Industrial Electronics Society, 2010, pp. 469-474

[11] H. Soleimani, A. Dastfan and A. Hajizade, "DC bus voltage control by using a multi input/output system of photovoltaic/battery," 2012 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES), 2012, pp. 1-5

[12] H. J. Bergveld, W. S. Kruijt, and P. H. L. Notten, “Basic information on batteries,” in Battery Management Systems Design by Modelling, Kluwer Academic Publishers, 2002, vol. 1, pp. 31–53

[13] R. W. Erickson and D. Maksimovic, Fundamentals of Power Electronics, Nowell, MA: Kluwer, 2001

[14] Slobodan N. Vukosavić, Grid-Side Converters Control and Design, Springer International Publishing, 2018

[15] A. S. Ahmed, B. A. Abdullah and W. G. A. Abdelaal, "MPPT algorithms: Performance and evaluation," 2016 11th International Conference on Computer Engineering & Systems (ICCES), 2016, pp. 461-467

[16] Y. -s. Bai and C. -n. Zhang, "Experiments study on fast charge technology for Lithium-ion electric vehicle batteries," 2014 IEEE Conference and Expo Transportation Electrification Asia-Pacific (ITEC Asia-Pacific), 2014, pp. 1-6

[17] C. Lin, C. Hsieh and K. Chen, "A Li-Ion Battery Charger With Smooth Control Circuit and Built-In Resistance Compensator for Achieving Stable and Fast Charging," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 57, no. 2, pp. 506-517, Feb. 2010

[18] M. Chen and G. Rincon-Mora, “Accurate, compact, and power-efficient Li-ion battery charger circuit,” IEEE Trans. Circuits Syst. II, Exp. Briefs, vol. 53, no. 11, pp. 1180–1184, Nov. 2006

[19] Y.-H. Liu, J.-H. Teng, and Y.-C. Lin, “Search for an optimal rapid charging pattern for lithium-ion batteries using ant colony system algorithm,” IEEE Trans. Ind. Electron., vol. 52, no. 5, pp. 1328–1336, Oct. 2005

[20] M. Vidyasagar, Nonlinear Systems Analysis, Society for Industrial and Applied Mathematics, 2002

ЕЛЕКТРИЧНА КОЛА, ЕЛЕКТРИЧНИ СИСТЕМИ И ОБРАДА СИГНАЛА
/
ELECTRIC CIRCUITS AND SYSTEMS, AND SIGNAL PROCESSING

(ЕК/EKI)

Kristalni filtri za opseg frekvencija 150-170MHz

Dragi Dujković, Irini Reljin, Lenkica Grubišić, Snežana Dedić-Nešić, Ana Gavrovska

***Apstrakt*—Tehnologija izrade filtara zasnovanih na kristalima kvarca je i dalje zahtevan tehnološki postupak iako je imao svoj maksimum osamdesetih godina prošlog veka. Međutim, i danas se koriste ove komponente u savremenim telekomunikacijama, naročito digitalnim, i to tamo gde je veoma bitno imati kvalitetne uređaje za prijem, predaju i prenos signala. Uslovi rada filtara i uređaja u koji su ugrađeni, pored električnih zahtevaju i dobre klimomehaničke radne karakteristike koje su definisane zahtevima korisnika. U radu su opisani kristalni filtri propusnici opsega, namenjeni za rad u spoljnim uslovima pri visokim frekvencijama, i to u opsegu od 150 do 170 MHz.**

***Ključne reči*—Kristali, filtriranje, projektovanje, kućište, tehničke karakteristike.**

I. Uvod

Kristalni filtri koji se implementiraju u uređaje za razne civilne i vojne namene, često imaju strogo definisane zahteve vezane za električne karakteristike, ali i za takozvane klimomehaničke karakteristike pri kojima su iskazani uslovi funkcionisanja filtra pri specifičnim uslovima namene (ovo je uobičajeno za uređaje koji se koriste van zaštite prostorija, tj. u spoljnim i terenskim uslovima rada).

U radu je opisan razvoj novih tipova kristalnih filtara propusnika opsega učestanosti na frekvencijama od 150 do 170MHz prema specifikacijama definisanim od strane korisnika. Specifikacija filtra obuhvata električne karakteristike koje definišu tok amplitudske karakteristike i klimomehaničke karakteristike koje definišu zahteve u pogledu vibracija, udara i temperature za rad filtra u specifičnim uslovima.

Razmatrani su izazovi ostvarivanja što većeg slabljenja u nepropusnom opsegu filtra i što manjeg slabljenja u propusnom opsegu filtra. Paralelno sa rešavanjem ovih problema odvijao se razvoj kristalnih jedinki u cilju ostvarivanja svih parametara koje je projekat filtra odredio a posebno ostvarivanja njihovog što većeg *Q*-faktora i što većeg faktora potiskivanja neželjenih rezonancija.

Takođe je urađena i analiza svih elemenata koji se u filtar ugrađuju i ispitan njihov uticaj na osetljivost filtra u pogledu funkcionisanja u radnom temperaturnom opsegu i pri ekstremnim zadatim uslovima rada [1-4].

Rad je organizovan na sledeći način. U drugoj glavi dat je kratak osvrt na potrebe za realizacijom ovakvih filtara. Treća glava je posvećena projektovanju filtara. Kristalne jedinke su opisane, kao i sama realizacija u glavi četiri. U petom poglavlju se nalazi zaključak.

Dragi Dujković – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar Kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: dragi@etf.bg.ac.rs).

Irini Reljin – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar Kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: dragi@etf.bg.ac.rs).

Lenkica Grubišić – Institut Mihajlo Pupin, Volgina 15, 11060 Beograd, Srbija (e-mail: lenka@piezo.imp.bg.ac.rs).

Snežana Dedić-Nešić– Institut Mihajlo Pupin, Volgina 15, 11060 Beograd, Srbija (e-mail: snezanadn@gmail.com).

Ana Gavrovska – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar Kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: dragi@etf.bg.ac.rs).

II. Potreba za realizacijom filtara konkretnih karakteristika

Zahtevane električne karakteristike filtara su direktno vezane za namenu uređaja, a ta namena određuje tok amplitudske i fazne karakteristike u propusnom opsegu, centralnu frekvenciju filtra, širinu propusnog opsega i selektivnost.

U zavisnosti od toga u kakvim će uslovima uređaj funkcionisati zahtev za specifikaciju pored standardnih električnih karakterisrtika ima i zahteve koji se odnose na vibracije, udare, potrese, temperaturu, pritisak, vlagu (tzv. klimomehaničke uslove rada) ali i mnoge druge specifične zahteve.

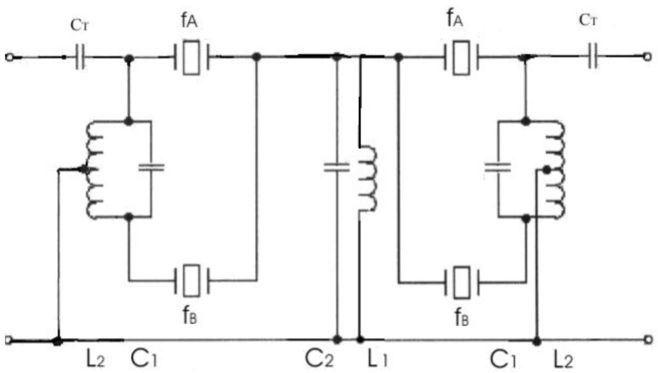
Ovakvi kristalni filtri za specijalne namene, koji se prema zadatoj specifikaciji ne mogu pronaći u katalozima proizvođača, zahtevaju razvoj uz odgovarajući projekat filtra prema postavljenim zahtevima.

U celom svetu postoji samo nekoliko proizvođača koji mogu realizovati kristalne filtre na osnovu specifičnih zahteva korisnika. To zahteva razvoj novih tehnoloških postupaka pri proizvodnji, kao i projektovanje filtara i njihovih komponenata. Zbog toga se većina proizvođača bavi proizvodnjom kataloških tipova filtara sa standardnim karakteristikama i komponentama [4–10].

III. Projektovanje filtara

Urađen je projekat filtra čija je lista tehničkih zahteva data u tabeli1. Na osnovu definisanih zahteva u pogledu toka amplitudske karakteristike u propusnom i nepropusnom opsegu određen je red filtra i mreža koja obezbeđuje ispunjavanje postavljenih zahteva (videti sliku 1). Pri proračunu filtra se mora voditi računa o tolerancijama elemenata koji se ugrađuju u filtar, da bi on mogao ispuniti tražene karakteristike.

Prema predviđenim gubicima u mreži, urađen je proračun filtra, određena električna šema i definisani zahtevi vezani za kristalne jedinke. Projektom filtra su tačno definisani svi zahtevi koji se odnose na parametre kristalnih jedinki, njihove vrednosti i tolerancije. Definisani su i zahtevi u pogledu frekvencija kristala, serijske i paralelne kapacitivnosti kristala i podešenosti frekvencije na sobnoj temperaturi. Takođe su definisane i maksimalna dozvoljena odstupanja frekvencija kristala u radnom temperaturnom opsegu filtra. Na osnovu zahteva za ispunjavanje klimomehaničkih uslova rada odabran je i tip kućišta kristalnih jedinki.



Sl. 1. Električna šema kristalnog filtra 156.800 MHz.

Prema definisanim zahtevima sadržanim u specifikaciji filtra (Tabela I) urađen je projekat filtra 4-og reda uz Chebyshev aproksimaciju. Razvijene su mikrominijaturne kristalne jedinke AT-reza sa malim odstupanjem frekvencije u širokom temperaturnom opsegu koje ispunjavaju postavljene zahteve u pogledu položaja i potisnutosti sporednih rezonancija i u pogledu zahtevanih klimo-mehaničkih karakteristika.

Proračunom filtra osim kristalnih jedinki određuju se i vrednosti ostalih komponenata, kao što su kalemovi i kondenzatori kao i njihov raspored na štampanoj ploči.

Treba imati u vidu da filtri rade na visokim frekvencijama gde prisustvo parazitnih kapacitivnosti značajno utiče na karakteristike filtara, tako da je raspored komponenata od izuzetnog značaja u rešavanju problema selektivnosti i slabljenja u nepropusnom opsegu filtra [1-5].

Da bi izmerili amplitudne i fazne karakteristike, za merenja je korišćen realizovani kristalni filtar centralne frekvencije 156.800 MHz. Ovaj filtar ima relativno slabljenje u propusnom opsegu +/- 7.5 kHz manje od 3dB. Izvan frekventnog opsega +/-60 kHz slabljenje je veće od 40 dB . U širem opsegu frekvencija, do +/- 50 MHz, relativno slabljenje je veće od 40 dB.

Minimalno pogonsko slabljenje filtra je manje od 6 dB. Talasnost u propusnom opsegu je manja od 1 dB. Ulazna i izlazna otpornost filtra iznose 50 Ω.

Radni temperaturni opseg filtra je -20º C do +70 ºC. Temperaturni opseg skladištenja (opseg temperatura u kojima filtar neće promeniti karakteristike dok se čuva u skladištu) je -40 ºC do +85 ºC. Filtar je smešten u kućište dimenzija (61×26.2×26.2) mm.

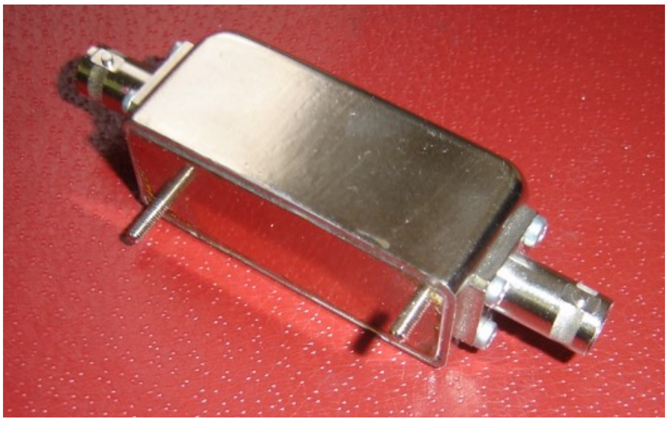
Filtar takođe mora da ispuni posebne zahteve vezane za vibracije, udare i temperaturu, tj. da zadovolji zahteve za klimomehaničke karakteristike rada. Detalji svih tehničkih karakteristika filtra su dati u Tabeli I.

IV. PRISTUP ZASNOVAN NA KRISTALNIM JEDINKAMA I REALIZOVANI KRISTALNI FILTAR

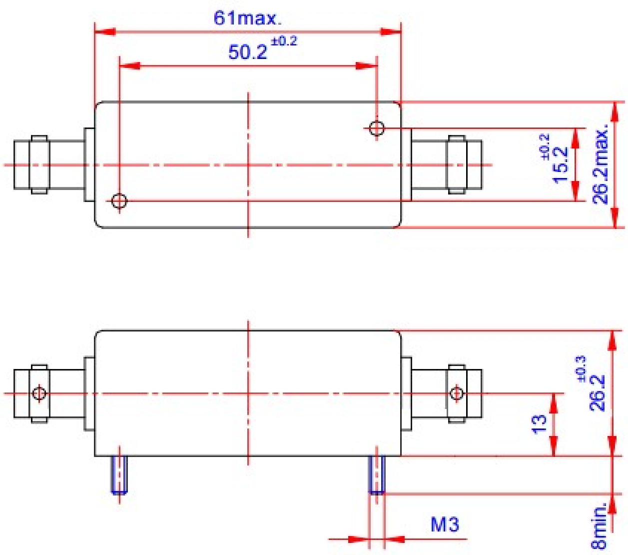
Kristalne jedinke (KJ) su planparalelne, prečnika Φ=5 mm. Upotrebljena je elektroda d=0.79mm. Planparalelne KJ su sečene u obliku kruga tako da su im bočne strane paralelne [10]. Kao elektrodni materijal korišćen je aluminijum. U proizvodnji se mora voditi računa o planparalelnosti radi ostvarenja zahteva vezanih za položaj i potisnutost sporednih rezonancija.

Za nosač kristalne jedinke upotrebljen je držač sa petljicama da bi se obezbedili najpovoljniji rezultati vezani za zahteve na otpornost rada kristalne jedinke pri uticaju vibracija i mehaničkih udara.

Na slici 2 prikazan je realizovani filtar. Filtar je smešten u standardno G10BNC kućište koje je prikazano na slici 3.



Sl. 2. Realizovani kristalni filtar.



Sl. 3. Kućište G10 BNC.

Rezultati ispitivanja filtara na sobnoj temperaturi i u radnom temperaturnom opsegu od -40°C do +85°C pokazuju da filtri ispunjavaju zahteve navedene u listi tehničkih podataka.

Rezultati merenja amplitudske karakteristike filtra na sobnoj temperaturi prikazani su u dijagramima na slici 4.

Na filtrima je sprovedeno merenje električnih karakteristika i testiranje prema zadatim uslovima. Konačni rezultati merenja pokazali su da su karakteristike realizovanih filtara u skladu sa zahtevima definisanim listom tehničkih podataka i da primenjeni tehnološki postupak daje visokokvalitetan filtar koji se može koristiti i u posebnim klimatskim i mehaničkim uslovima [4-8].

Lista tehničkih podataka za kristalne filtre 150-170 MHz je sledeća:

1. Kućište CW-HC-45
2. Q-faktor > 60000
3. dinamička kapacitivnost $C_1 = 0.85$ mpF ± 10%
4. paralelna kapacitivnost $C_o = 0.6$ pF ± 5%
5. dinamička otpornost $R_1 < 220\ \Omega$
6. podešenost df/f=±5 ppm
7. odstupanje df/f=±20 ppm
8. starenje df/f=1 ppm/god
9. radni temperaturni opseg -20÷+70˚C
10. neželjene rezonancije A. f_0+50 kHz–bez f_n
    B. f_n>30 dB za f_0±1 MHz

Tabela I Tehničke karakteristike filtara

Centralna frekvencija(CF)	150 - 170 MHz
Širina propusnog opsega na 3 dB	+/- 7.5 kHz
Talasnost u propusnom opsegu	1 dB max u opsegu +/- 7.5 kHz
Širina nepropusnog opsega na 40 dB	+/- 60 kHz max
Relativno slabljenje u nepropusnom opsegu	40 dB min za +/- 50 MHz
Minimalno pogonsko slabljenje	6 dB max
Ulazna impedansa	50 Ω
Izlazna impedansa	50 Ω
Dozvoljeni nivo neželjenih rezonancija	10 dB min
Radni temperaturni opseg	-20 ºC до +70 ºC
Temperaturni opseg skladištenja	-40 ºC до +85 ºC
Vibracije sinusne	10 do 2000 Hz 30 g
Udari	100 g 6 ms
Termički šok 5 ciklusa	-55 ºC do +125 ºC

V. ZAKLJUČAK

Pored proračuna i realizacije kristalnog filtra, razvijena je i nova kristalna jedinka. Ovakvi kristalni filtri predstavljaju novi proizvod, jer su u njemu korišćene nove tehnologije i nove komponente. Ovi proizvodi imaju širok dijapazon primena i veliku upotrebnu vrednost i ističu se svojom cenom i svojim kvalitetom, tako da su konkurentni ne samo na domaćem tržištu već i u celom svetu. Navedene karakteristike i konkurentnost na svetskom tržištu daju perspektivu razvoju novih elektronskih sklopova i proizvoda na bazi kristalne jedinke.

U daljem razvoju ovih i sličnih uređaja treba ići na usvajanje novih tehnologija izrade kristalnih jedinki i upotrebe novih i kvalitetnijih komponenti u kolu elektronskih sklopova novih uređaja [4-8].

ZAHVALNICA

Istraživanja opisana u ovom radu su delimično finansirana od strane Ministarstva za obrazovanje i nauku Republike Srbije.

LITERATURA

[1] Snežana Dedić-Nešić “Prilog projektovanju kristalnih filtara sa linearnom faznom karakteristikom” Magistarska teza, Elektrotehnički fakultet Beograd, 1991.

[2] A.I.Zverev: “Handbook of Filter Synthesis” John Wiley and Sons, Inc, 1967

[3] H. J. Blinchikoff, A. I. Zverev “Filtering in the Time and Frequency Domains” John Wiley and Sons, Inc, 1976

[4] D. S. Humpherys “The Analysis, Design and Synthesis of Electrical Filters” Prentice Hall, Englewood Clifs, N.J. 1970

[5] Snežana Dedić-Nešić, Dragi Dujković, Lenkica Grubišić, Irini Reljin, Branimir Reljin “Novi visokokvalitetni 90.1105 MHz filtri za komunikacije u rudnicima” ETRAN, 2010

[6] Dragi Dujković, Branimir Reljin, Snežana Dedić-Nešić, Lenkica Grubišić, Dubravka Jevtić , “Novi kristalnih filtar F121” ETRAN, 2012

[7] D. M. Dujković, S. Dedić–Nešić , L. Grubišić, A. Gavrovska, and I. Reljin, “High-quality crystal filter: Design and realization,” Telfor J., vol. 5, no. 2, pp. 118–122, 2013.

[8] D. M. Dujković, S. Dedić–Nešić, L. Grubišić, M. Paskaš, and I.Reljin, “Matched Crystal Filters” Telfor Proceedings, pp. pp. 481-484, Belgrade 2014.

[9] D. M. Dujković, S. Dedić–Nešić, L. Grubišić, M. Paskaš, and I. Reljin, “A Novel 76.8375 MHz Crystal Filter Based on Third Overtone Crystal Units“, Telfor Proceedings, pp. 177-181, Belgrade 2020.

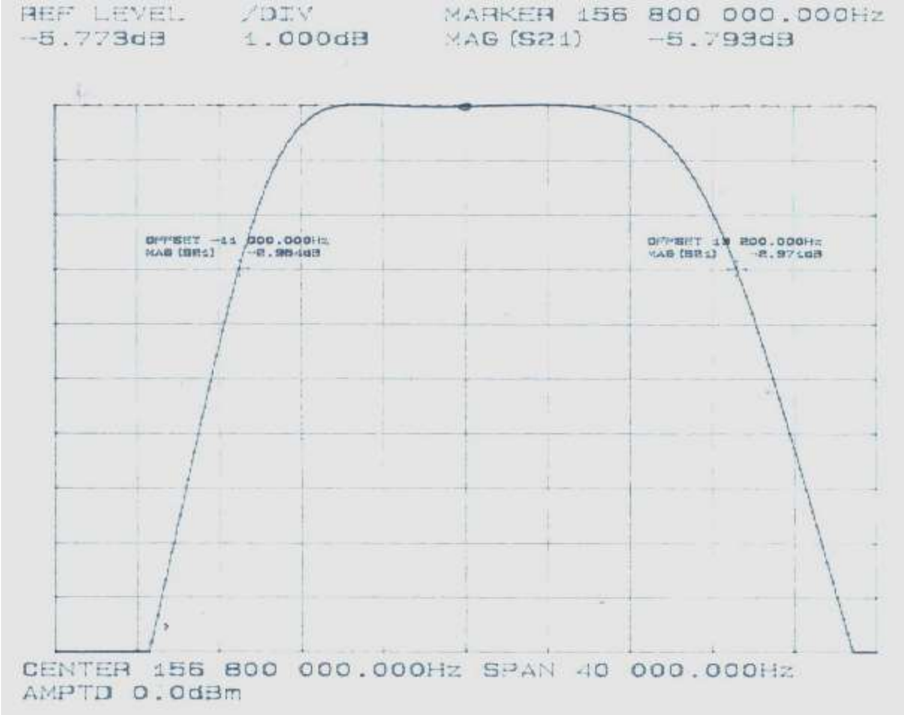
[10] L. Grubišić, S. Dedić–Nešić, D. M. Dujković, , B. Reljin, “Novi tip kristalne jedinke za filtre namenjene za rad u posebnim klimo mehaničkim uslovima “, ETRAN, 2012, EK2.6-1-4, Banja Vrućica 2011.

ABSTRACT

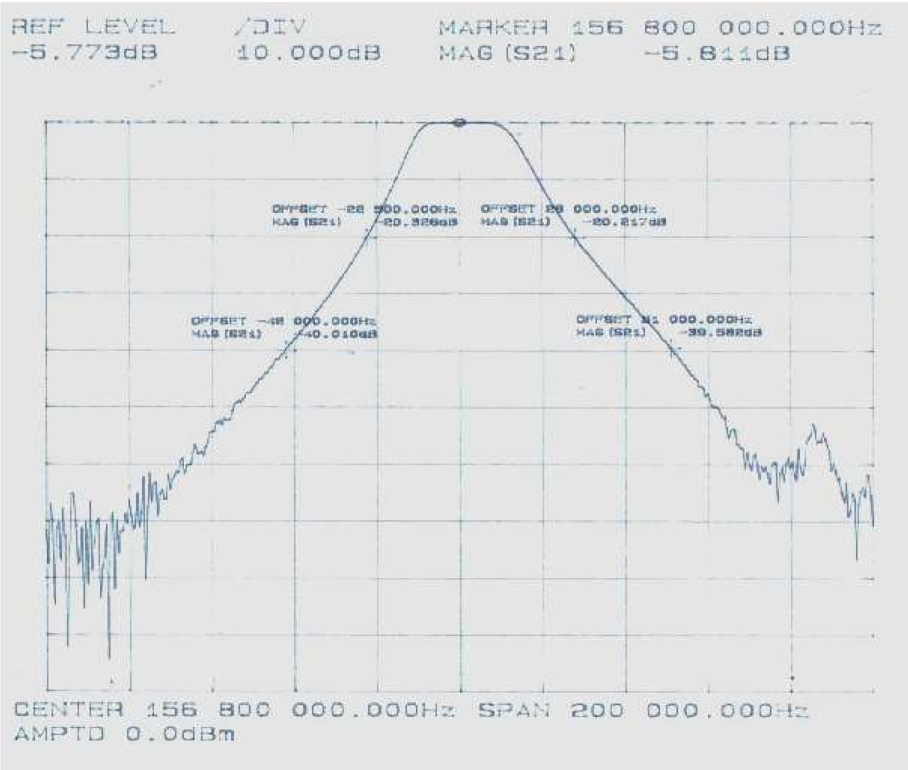
The technology of making filters based on quartz crystals is still a demanding technological process, although it had reached its maximum in the 1980s. However, even today, these components are used in modern telecommunications, especially digital, where it is very important to have high quality devices for receiving, transmitting and signal distribution. The operating conditions of the filters and devices in which they are installed, in addition to electrical ones, require good temperature-mechanical operating characteristics that are defined by the requirements made by users. The paper describes band-pass crystal filters, which are intended for operating in outdoor conditions at high frequencies in the range from 150 to 170MHz.

Crystal filters for frequency range 150-170 MHz

Dragi Dujkovic, Irini Reljin, Lenkica Grubisic, Snezana Dedic-Nesic, Ana Gavrovska



a)



b)

Sl. 4. Rezultati merenja amplitudske karakteristike filtra 156.800 MHz na sobnoj temperaturi, a) u okviru radnog opsega +/- 7.5 kHz od CF , b) u okviru opsega od +/- 50 MHz od CF

Covid-19 and other CT Scan Authentication using Wavelet based Watermarking

Amra Gicić, and Ana Gavrovska, *Member, IEEE*

Abstract— **Nowadays, it is essential to ensure the integrity of the medical image, especially for adequate region of interest (ROI) before taking any diagnostic decision, where watermarking is often used. In this paper, Covid-19 and other CT (Computerized Tomography) scans are analyzed using wavelet based watermarking under JPEG (Joint Picture Experts Group) compression attack. The structured method consisted of robust and fragile watermark, having in mind ROI, is implemented. JPEG compression attack is chosen since it is often used while saving case studies. This is of particular importance in the case of widely available Covid-19 CT scans of different resolutions. The CT scan structured watermarking shows promising results under JPEG attack. The results seem promising in both detection of image manipulation through fragile watermark and integrating visual logos through more robust approach.**

Index Terms— **Medical images, Computerized Tomography, watermarking, wavelet, JPEG, DICOM, ROI.**

I. Introduction

Security risks of medical images can vary from unauthorized access to errors occurring during transmission in hospital information systems (HIS) and PACS (Picture archiving and communication system) [1-4]. An attached file or a packet header (as in DICOM (Digital Imaging and Communications in Medicine) files) often carries all information needed to identify a particular image. However, keeping image metadata in a separate header file is prone to forgeries or clumsy practices. An alternative would be to embed all such information into the image content itself.

Medical data security has become inevitable in smart hospital applications to ensure data privacy and patient data security of Electronic Patient/Health Records (EPR/EHR) [5-7]. Medical reports and images are transferred to physicians at distant locations and to other hospitals for diagnosis. Before making any diagnostic decision, the integrity of region of interest (ROI) of the received medical image must be verified. Watermarking can be used to ensure integrity and authentication of the medical image. Main challenges associated with e-healthcare systems and digital watermarking are following:

Amra Gicić is with the School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11020 Belgrade, Serbia (e-mail: amragicic1625@gmail.com).

Ana Gavrovska is with the School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11020 Belgrade, Serbia (e-mails: anaga777@gmail.com; anaga777@etf.rs).

• medical image and EPR transmission should not cause separation between metadata and image content,

• data (visual and non-visual) should not be modified accidentally or intentionally while transmitting over the insecure medium,

• medical information authentication should be ensured to obtain confirmation of content that belongs to the appropriate patient (origin and integrity authentication).

In this paper analysis is performed based on a structured watermarking method in order to protect medical images (modalities). The quality of the medical image, such as CT (computed/computerized tomography) scan, is considered for further use having in mind common JPEG (Joint Picture Experts Group) compression attack. Furthermore, a comparison is made with another structured watermarking method to verify the benefit of the implemented model.

The paper is organized as follows. After the introduction, in Section II ROI and RONI (region of non-interest) based watermarking is explained, as well as common techniques used for medical image watermarking. Implementation and extraction of the proposed structured RONI based watermark are briefly given in Section III. Experimental results are presented in Section IV. Conclusion is given in Section V.

II. ROI and RONI based Watermarking for Medical Images

A. *Medical image watermarking*

There are a lot of different watermarking methods, where wavelet and LSB (Least Significant Bit) based approaches are one of the most common [8-16]. Coatrieux et al. [8] proposed Region of Interest (ROI) based approach to preserve the diagnostically relevant region, and Region of Non-Interest (RONI) usage in order to keep integrity and to serve as watermark carrier. Mehta et al. [9] studied the performance of three different techniques in watermarking: DWT (Discrete Wavelet Transform), SVD (Singular Value Decomposition) and hybrid (DWT-SVD) based watermarking. Similarly, Navas et al. [10] proposed a method of non-blind transform domain watermarking using DWT-DCT-SVD approach (DCT- Discrete Cosines Transform). Fragile watermarking methods can be focused on ROI integrity, where approaches can also be block-based (Liew et al. [11], Tjokorda Agung [12]). Also, watermarking methods are designed for various modalities (e.g. Nambakhsh et al. [13] for protecting positron emission tomography - PET images, Castiglione et al. [14] for functional magnetic resonance imaging - fMRI images). Giakoumaki et al. [15] discussed the watermarking perspectives in a range of medical data management and distribution. The most common metadata structure correspond to DICOM standard, where the data is saved in the image header as a part of the image file and includes information related to patient, hospital, image, acquisition properties [16].

Different algorithms are available for ROI segmentation, and this depends on the medical image type. A physician can choose ROI in an image, where ROI represents rectangular area. In Fig.1 an example of segmentation of the ROI part for a CT image using a rectangular selection (rectangle) is illustrated.

Fig. 1. ROI segmentation for a CT image.

Moreover, there are various watermarks available. In Fig.2 visual and textual watermarks are illustrated: potential logo (logotype) representing clinics and text (txt) file representing information about the patient (PatientInfo), information about the image (ImageInfo), data about the diagnosis (DiagnosisInfo), data about the physician (PhysicianInfo). Such and additional data can also be found in the headers of DICOM format.

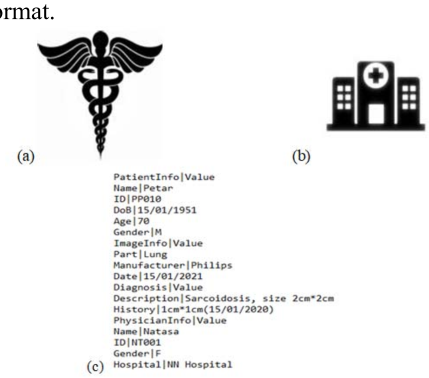


Fig. 2. Examples of visual watermarks (logos): (a) Logo1, (b) Logo2, and (c) a textual watermark.

B. *Robust and Fragile medical image authentication*

Watermarking often tends to be invisible and robust. On the other hand, some watermarks can be considered fragile if their usage is needed for detection of content manipulation. Thus, watermark can be considered as multiple or structured, having both robust and fragile part. If content manipulation occurs, fragile approach should prevent fragile watermark extraction. Fragile watermark can be of textual type, representing metadata. Also, robust watermark like logo can be invisible.

DWT enables energy compact representation of an image [9-10]. Using 1-level DWT decomposition image is transformed into four sub-bands: low-high (LH(1)), high-low (HL(1)), high-high (HH(1)) and low-low (LL(1)) band. LL band represents low frequency, HL and LH represent middle frequency (HL - horizontal and LH - vertical details) and HH represents high frequency band (diagonal image details), respectively. In 2-level DWT decomposition LL from the first level (noted as LL1) is further decomposed to additional four bands (LL2, LH2, HL2, HH2). The bands represented by wavelet coefficients can be used for embedding fragile and robust watermarks.

Direct embedding on the wavelet coefficients is not necessary. The embedding can be performed on the elements of singular values of the DWT sub-bands. This is done by SVD, where SVD decomposes a matrix into three sub-matrices in such a way that singular values get separated in the form of diagonal matrix [9]. The three decomposed matrices are: left singular matrix *U*, singular matrix *S*, and right singular matrix *V*, where *U* and *V* are the unitary matrices, and *S* is a diagonal matrix. Also, hidden information can be stored into specific LSB positions depending on the secret key. A sparse matrix generated by a private-key to determine the location of non-ROI modification, enhance the security of patient data [17-18].

In this paper a multiple medical image watermarking scheme based on DWT and SVD is implemented for robust watermark, and LSB based fragile watermarking technique for tamper detection using BCH *(Bose-Chaudhuri-Hocquenghem)* coding for noise detection and correction to enhance protection. The implementation of the BCH requires choosing a polynomial called the generator that is known to the transmitter and the receiver [19]. The transmitter performs the encoding procedure on the message stream to generate a certain number of check bits called a checksum. This checksum is appended to the message being transmitted. The receiver performs the decoding procedure, to verify that the checksum is valid. A BCH code with symbols from Galois field *GF*(2) (with two elements 0 and 1), with length 31 of the code word (message + checksum), and 16 message length which represents the binary message are used.

III. EXPERIMENTAL ANALYSIS

A. *Implementation of the structured RONI based watermark*

In Fig.3 block diagram for implementation of the structured watermark is illustrated.

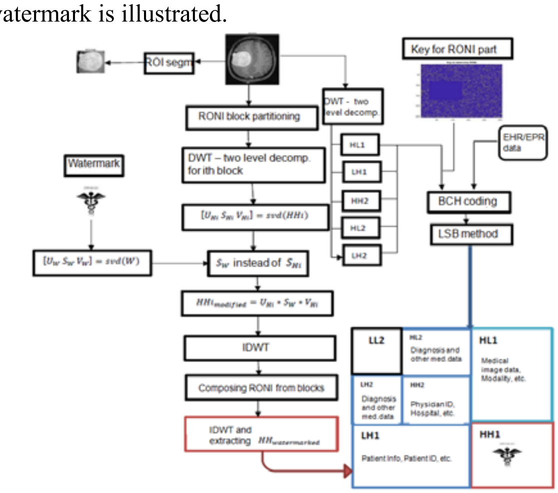


Fig. 3. Block diagram for implementation of the structured RONI and wavelet based watermark.

Robust watermark is implemented in transform domain, where 2-level DWT decomposition is applied, as well as SVD. ROI segmentation is performed, where its limits are memorized. This is followed by RONI partitioning to blocks of 8x8 pixels. For each block 2-level DWT is applied using Haar wavelet obtaining components noted as LL2, HL2, LH2, HH2, HL1, LH1 i HH1. Then, SVD is applied to HH1. SVD components of (logo) watermark are integrated to SVD components of the image. Finally, IDWT is applied for all the blocks of RONI part in order to compose the watermarked RONI. Watermarked HH1 part is related to robust approach.

For the fragile watermark, 2-level DWT is applied using Haar wavelet. Additional data is implemented in all parts except in the lowest and the highest frequencies (LL1, HH1). Textual data is divided into remaining bands, so that the diagnosis data belongs to HL2 and LH2, physician and hospital data belongs to HH2, patient info data belongs to LH1, and image data belongs to HL1 part. Private Key is generated in order not to affect the ROI while LSB method is applied for fragile watermark. This is a matrix of 256x256 pixels consisted of random numbers from 0 to 10, where zeros become 1, and the rest of the values become 0. Length of the message for watermarking is defined and the message itself is converted to binary representation. LSB method is applied along with BCH coding for textual data integration and for improved data security in order to keep data safe from unauthorized access. Coding word is of 31 bit length, while the message length is 16 bits, and the control sum is checked. If image pixel's LSB, where watermark is embedded equals to a corresponding bit of hidden message, the pixel remains the same. If this is not the case, bit corresponding to the message is put in that place. Finally, IDWT is applied on watermarked image so that fragile and robust parts can be composed into one final image.

B. Digital watermark extraction

Digital watermark extraction is illustrated in Fig.4.

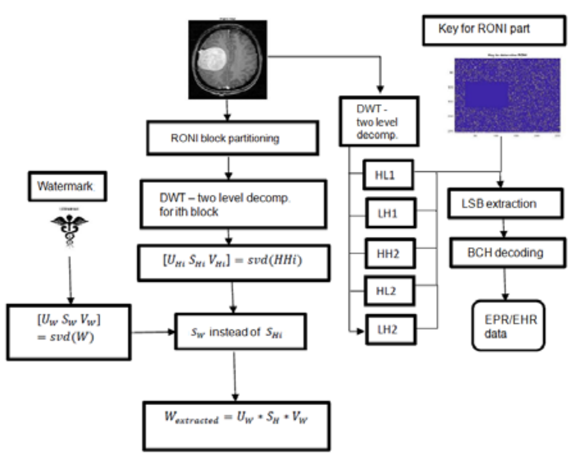


Fig. 4. Block diagram for digital watermark extraction.

Robust watermark is extracted similarly as in the implementation step. The ROI limits are used for obtaining RONI part of the watermarked image. This is followed by its partitioning, and applying 2-level DWT for obtaining the common ranges. The SVD approach is applied for HH1 part, and IDWT for logo watermark extraction. Similarly, as in the previous step with LSB extraction and BCH decoding fragile medical data can be extracted.

C. Datasets and experimental phases

Two sets are tested. The examples from the first set are presented in Fig.5 (PNG1-5), where the CT scans from [20] represent 24bit Covid-19 images in png format. Also, cancer CT [21], which are 16bit of the same resolution and in DICOM format, are investigated. They are illustrated in Fig.6 (DIC1-5). Here, only rectangle shape is assumed for ROI. ROI is selected manually, and is not part of this research. Thus, the method can be considered semi-automatic.

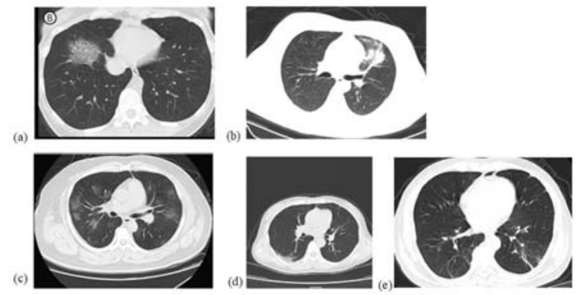


Fig. 5. Covid-19 CT chest scans: (a) 406x302 image, (b) 320x430 image, (c) 406x304 image, (d) 501x374 image, (e) 589x448 image, example.

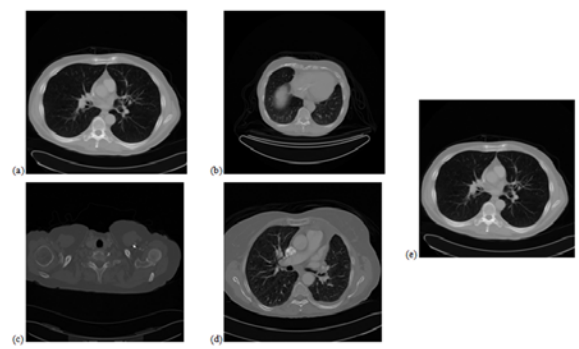


Fig. 6. Cancer CT 512x512 scans: (a)-(d) representing five DICOM images.

After the structured watermark integration and extraction model implementation, two logo watermarks (128x128) from Fig.2(a)-(b), and appropriate medical textual data as in Fig.2(c) are tested having in mind JPEG attack, as a common attack in image forensics. In the first phase of experimental analysis, the Covid-19 scans are tested, where PSNR (Peak-to-Signal Noise Ratio) and SSIM (Structural Similarity Index Metric) are calculated:

$$PSNR_{db}\left(I_{ref}, I_{tst}\right) = 10\log_{10}\frac{MAX_I^2}{MSE}, \quad (1)$$

$$MSE = \frac{1}{NxM}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}\left(I_{ref}(i,j) - I_{tst}(i,j)\right)^2, \quad (2)$$

$$SSIM\left(I_{ref}, I_{tst}\right) = L(I_{ref}, I_{tst})C(I_{ref}, I_{tst})S(I_{ref}, I_{tst}) \quad (3)$$

where in (1) MAX_I represents maximum intensity value and MSE is explained in (2). In (3) three factors representing luminance, contrast, and structure (L, C, S, respectively) are used for comparing reference and tested image (I_{ref}, I_{tst}). The second phase is similar, and PSNR and SSIM are calculated for the cancer CT dicom images, having in mind JPEG attack. Four different quality JPEG levels are tested (Q = 15, 30, 75, 100). In the third phase a comparison to [22] is performed.

IV. Experimental results

A. *Experimental results for Covid-19 CT data under JPEG attacks*

The obtained PSNR and SSIM results for Logo1 and Logo2, are shown in Fig.7(a) and Fig.7(b), respectively. The results are obtained for four JPEG quality (Q) levels, and five images PNG1-5. Extracted watermark Logo1 is shown in Fig.8. Similarly, the extraction results for Logo2 are shown in Fig.9.

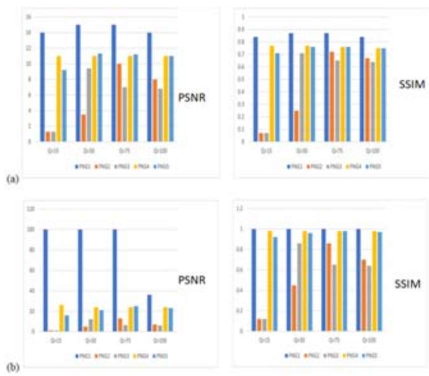


Fig. 7. PSNR and SSIM results under JPEG attacks for Covid-19 CT scans, and: (a) Logo1 and (b) Logo2.

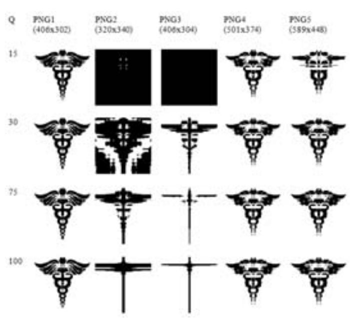


Fig. 8. Logo1 extraction results for Covid-19 scans under JPEG attacks.

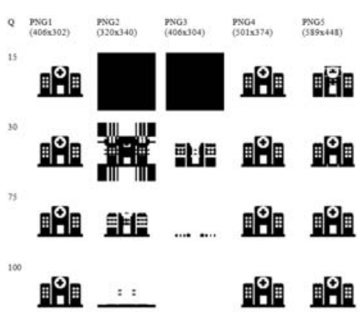


Fig. 9. Logo2 extraction results for Covid-19 scans under JPEG attacks.

B. *Experimental results for Cancer CT data under JPEG attacks*

The obtained PSNR and SSIM results for Logo1 and Logo2, are shown in Fig.10(a) and Fig.10(b), respectively. Extracted logos example is presented in Fig.11.

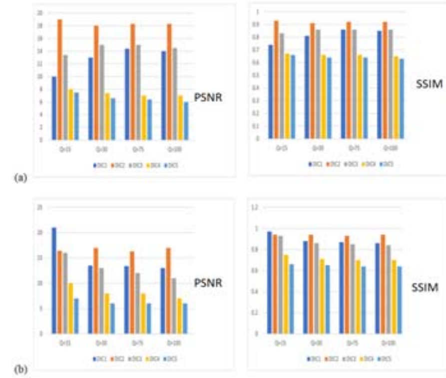


Fig. 10. PSNR and SSIM results under JPEG attacks for cancer CT scans, and: (a) Logo1 and (b) Logo2.

Fig. 11. Extracted (a) Logo1 and (b) Logo2 for DIC1 image and Q=30 quality.

C. *Comparison results*

The comparison is performed with another structured approach from literature [22]. The comparison results for PNG1-5 are presented in Fig.12. Similarly, the results for DICOM images are given in Fig.13, where Q1 denotes the quality for the method from [22].

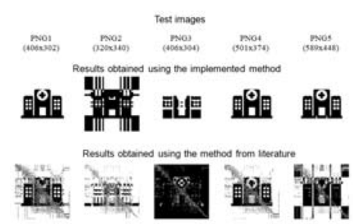


Fig. 12. Comparison results between the implemented method and the method from literature [22] for PNG images.

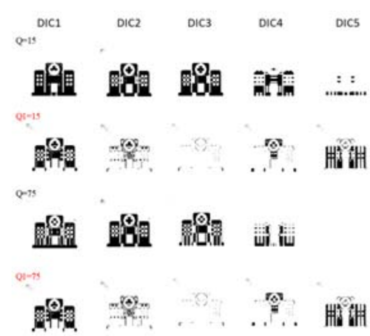


Fig. 13. Comparison results between the implemented method and the method from [22] for DICOM images.

PSNR and SSIM comparison results are illustrated in Fig.14 and Fig.15 for the first and the second dataset, respectively. The proposed method is compared to [22] which is based on DWT, Hessenberg decomposition and SVD.

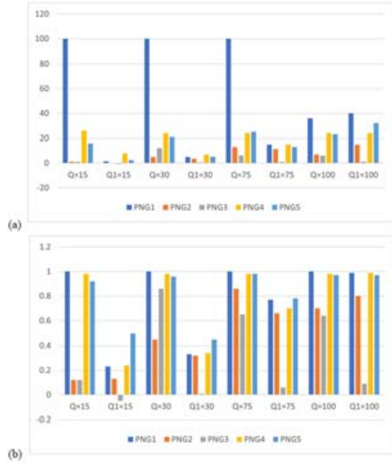


Fig. 14. (a) PSNR and (b) SSIM comparison results between the implemented method (Q) and the method [22] (Q1) for PNG images.

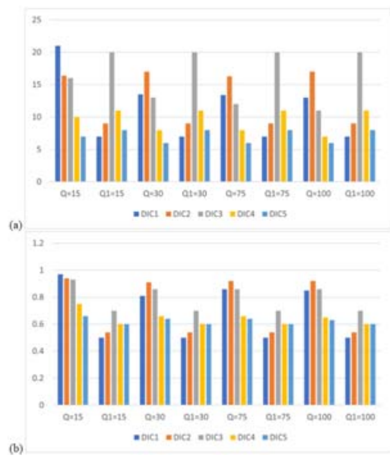


Fig. 15. (a) PSNR and (b) SSIM comparison results between the implemented method (Q) and the method [22] (Q1) for DICOM images.

In JPEG compression and similar tested attacks (except Gaussian filtering) fragile watermark is not extracted as original, which confirms that the manipulation occurred. Example of the extracted fragile watermark is shown in Fig.16.

Fig. 16 Example of unapropriatelly fragile watermark extraction.

JPEG compression could be recommended in certain cases, especially when medical or professional staff communicates via modern mobile applications that do not include support for the DICOM format (DICOM viewer). This should be done only for the purpose of more efficient informal communication.

In the current situation of the corona virus COVID-19, remote consultations were proposed to avoid cross-infection and regional differences in medical resources. However, the safety of digital medical imaging in remote consultations has also attracted more attention from the medical industry.

Covid-19 cases are often of different resolutions because of the use of different equipment and because of more efficient analyzes. With such data, ROI often has large dimensions, so it often occupies a larger part of the original image, which reduces the watermarking capacity. The influence of capacity on the robustness and perceptibility of watermarked image is not negligible. By increasing the data payload, the robustness will decrease and the perceptibility will increase. Dimensions (resolution) of the host image should also be further considered.

Image adjustment of size 512x512 pixels was applied before embedding the watermark. The worst results were obtained for PNG2 and PNG3, as shown for the extraction of both visual robust watermarks in Fig. 8-9. The PNG2 example has a lower resolution compared to the predefined image dimension adjustment, while the reason for PNG3 example was the image content itself, because similar image dimensions are in the PNG1 example. A similar experiment was done for dicom files, which correspond to cancer diagnostics. The obtained results are satisfactory because in most cases the watermark proved to be robust, i.e. it can be recognized. Text data during this attack in all formats could not be detected, which was the intention. The visual quality changes of selected medical images after watermark embedding are not noticeable, while the general image quality after watermark extraction is directly correlated with the extracted watermark quality discussed in this paper. Nevertheless, the RONI approach is recommended in order to prevent possible ROI changes.

The main contribution of this paper is the implementation of the DWT and RONI based watermarking for Covid-19 and other CT scans using robust and fragile approach, where the results showed the need for systematic approach for authentication in the case of available scans during pandemics, particularly having in mind JPEG attacks. The future experiments need to be performed for larger CT databases since other issues may occur besides lossy compression (e.g. due to segmentation).

V. CONCLUSION

This paper presents multiple/structured watermarking method which is based on RONI, as well a DWT, SVD, LSB and BCH techniques. Both robust and fragile watermarks are implemented. Analysis under JPEG compression attack showed watermark characteristics that can be useful for CT, like Covid-19 scans of different resolutions. Experimental results show that the model gives a good compromise between imperceptibility, robustness and fragility.

Future research should be directed towards enabling automatic CT scan watermarking, and further testing with larger database and watermarking method capacity.

Acknowledgment

This work was partially supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

References

[1] A. Anand and A. K. Singh, "Watermarking techniques for medical data authentication: a survey," *Multimedia Tools and Applications* pp. 1-33, 2020. https://doi.org/10.1007/s11042-020-08801-0.

[2] D. Laishram, and K. M. Singh, "A watermarking scheme for source authentication, ownership identification, tamper detection and restoration for color medical images," *Multimedia Tools and Applications*, pp. 1-61, 2021. https://doi.org/10.1007/s11042-020-10389-4

[3] X. Zhou, Y. Ma, Q. Zhang, M. A. Mohammed, and R. Damaševičius. "A Reversible Watermarking System for Medical Color Images: Balancing Capacity, Imperceptibility, and Robustness," *Electronics* 10, no. 9, pp. 1024, 2021.

[4] PACS sistemi, http://www.telemed.rs/pacs-sistemi/ (last accessed 03.05.2021.)

[5] A. Shehab, M. Elhoseny, K. Muhammad, A. K. Sangaiah, P. Yang, H. Huang, and G. Hou, "Secure and robust fragile watermarking scheme for medical images," *IEEE Access*, *6*, pp. 10269-10278, 2018.

[6] L. O. M. Kobayashi, S. S. Furuie, P. S. L.M. Barreto, "Providing integrity and authenticity in DICOM images: a novel approach," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 4, pp. 582-589, 2009.

[7] V. Diaz-Flores-Garcia, et al., "Detecting the manipulation of digital clinical records in dental practice," *Radiography*, 23.4, pp.e103-e107, 2017.

[8] G. Coatrieux, H. Maitre, B. Sankur, "Strict integrity control of biomedical images," In *Security and watermarking of multimedia contents III,* International Society for Optics and Photonics, pp. 229-240, 2001.

[9] S. Mehta, et al., "A study of DWT and SVD based watermarking algorithms for patient privacy in medical images," In *2013 IEEE International Conference on Healthcare Informatics*, IEEE, pp. 287-296. 2013.

[10] K. A. Navas, et al., "Dwt-dct-svd based watermarking," In 3rd Int. Conference on Communication Systems Software and Middleware and Workshops (COMSWARE'08), IEEE, pp. 271-274, 2008.

[11] S.C. Liew, S.W. Liew, and J.M. Zain, "Reversible medical image watermarking for tamper detection and recovery with Run Length Encoding compression," *World Academy of Science, Engineering and Technology*, 72, pp. 799-803, 2010.

[12] T. BW Agung, et al., "Medical image watermarking with tamper detection and recovery using reversible watermarking with LSB modification and run length encoding (RLE) compression," In *2012 IEEE International Conference on Communication, Networks and Satellite (ComNetSat)*, IEEE, pp. 167-171, 2012.

[13] M. S. Nambakhsh, A. Ahmadian, and H. Zaidi, "A contextual based double watermarking of PET images by patient ID and ECG signal," *Computer methods and programs in biomedicine*, 104.3, pp. 418-425, 2011.

[14] A. Castiglione, et al., "On the protection of fMRI images in multi-domain environments," In: *2015 IEEE 29th International Conference on Advanced Information Networking and Applications*, IEEE, pp. 476-481, 2015.

[15] A. Giakoumaki, S. Pavlopoulos, D. Koutsouris, "Multiple image watermarking applied to health information management," *IEEE Transactions on Information Technology in Biomedicine*, 10.4, pp. 722-732, 2006.

[16] O. S. Pianykh, *Digital imaging and communications in medicine (DICOM): a practical introduction and survival guide*, Vol. 6, Berlin: Springer, 2008.

[17] Github, https://github.com/Sherry-SR/Wavelets_Watermarking (last accessed 01.04.2021.)

[18] Github, https://github.com/ArkaPM/Region-of-Interest-Based-Medical-Image-Compression (last accessed 01.04.2021.)

[19] Y. S. Han, *BCH codes*, Graduate Institute of Communication Engineering, National Taipei University Taiwan, 2010. http://ct.ee.ntust.edu.tw/BCH_code.pdf (last accessed 04.05.2021.)

[20] CT Covid data, https://github.com/UCSD-AI4H/COVID-CT/blob/master/Images-processed/CT_COVID.zip (last accessed 31.5.2021.)

[21] CT Medical Images, https://www.kaggle.com/kmader/siim-medical-images?select=dicom_dir (last accessed 31.5.2021.)

[22] Github, https://github.com/Saeid-jhn/Digital-Image-Watermarking (last accessed 01.06.2021.)

ЕЛЕКТРОНИКА / ELECTRONICS

(ЕЛ/ELI)

Monitoring system for AC current up to 20A

Milan Savić, Dejan Stevanović and Miona Andrejević Stošović, *Member, IEEE*

***Abstract*—This paper presents a system for AC current monitoring in home appliances up to 20A. It is implemented on a custom made PCB. System also measures voltage, line frequency, power factor, active power and total imported active energy. Measurement results can be obtained by a remote computer or some other device via serial RS-485 interface. Consumer is enabled to have bigger control over real-time current consumption by installing several monitoring devices and connecting them into a network.**

***Index Terms*—Energy efficiency, Modbus RTU, monitoring system.**

I. Introduction

Energy efficiency has a fundamental role to play in the transition towards a more competitive, secure and sustainable energy system. Although energy powers our societies and economies, future growth must be driven with less energy and lower costs. According to the Energy Efficiency Communication of July 2014 [1], the EU is expected to achieve energy savings of 18%-19% by 2020. Reduced power consumption leads to reduced emissions and, consequently, reduced carbon footprint. This is a straightforward benefit. However, this could happen only if EU countries implement all of the existing legislation on energy efficiency. Unfortunately, efficiency of electrical distribution is currently not much managed or planned by utilities. The unfavourable result is that most utilities waste considerable amounts of electricity. For example, the annual value of transmission and distribution losses runs up to 6% of total generated energy [2]. These losses mainly occur in the low and medium voltage lines, and also in primary and secondary substations. One way for reducing losses and increasing efficiency on low level power distribution system is improving the system for registration of electric energy consumption [3-4]. Another way is to implement home energy monitoring system. A lot of similar systems have been already developed [5]. Some of realized systems are described in [6-8].

Economic return is one of the major reasons why households should consider and adopt smart energy management products. A home energy monitoring system allows consumers to have significant role in energy management activities. It can be implemented by using smart sockets. Probably, the simplest and most straightforward way to monitor and control energy usage is by replacing traditional sockets and plugs with the smart ones.

Milan Savić is with the College of Applied Technical Sciences Niš, Aleksandra Medvedeva 20, 18000 Niš, Republic of Serbia (e-mail: milan.savic@ vtsnis.edu.rs).

Dejan Stevanović and Miona Andrejević Stošović are with University of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, Niš, Serbia,{dejan.stevanovic, miona.andrejevic}@elfak.ni.ac.rs.

Namely, placing multiple measurement devices in the household enables that consumer has nearly instant information about current consumption of each appliance. Some smart sockets contain relay, so that turning load on and off is supported. In this paper we proposed a system that can be implemented at electric panel. Moreover, our system can be used as smart fuses.

This paper is organized as follows. In the next section, the basic definitions that correlate power parameters with measured current and voltage data will be given. The third chapter will be dedicated to description of realised system, while in the fourth chapter measurement results will be given. The conclusion is in the fifth section.

II. Definitions of electric power quantity

The core of our system is a MCP39F521 [9] which calculates all power quantities that are of interest for utility to control consumption. Usually, these values are defined by appropriate standards. All these circuits relay on digital signal processing of voltage and current samples. The instantaneous value of voltage and current are attenuated through voltage divider, while for current can be used current transformers, Rogowski coil sensors or shunt resistors. The first set of signal conditioning that occurs inside MCP39F521 is shown in Fig.1.

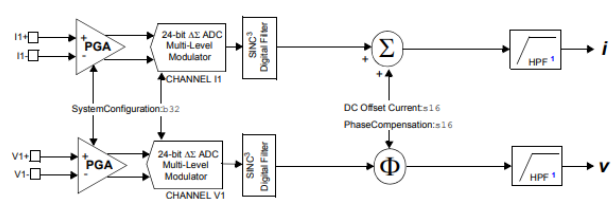


Fig. 1. Channel I1 and V1 signal flow

The obtained signal at output of attenuator is sent to ADC where it is sampled at discrete time points (at least two per period, according to the Nyquist-Shannon theorem) and digitalized. DSP processes digital voltage and current samples and calculates all necessary power quantities. Instantaneous value of signal (current or voltage) in time domain can be expressed as:

$$x(t) = \sqrt{2} X_{RMS} \cdot \cos(2\pi f t + \varphi) \qquad (1)$$

After the discretization in equidistant time intervals, it is transformed to:

$$x(nT) = \sqrt{2} X_{RMS} \cdot \cos(2\pi \frac{f}{f_{sempl}} n + \varphi) \quad , \qquad (2)$$

where f and f_{sempl}, are frequency of the signal and the sampling frequency, respectively. The RMS value is calculated using the following equation:

$$X_{RMS} = \sqrt{\frac{\sum_{n=1}^{N} x(nT)^2}{N}} \quad . \qquad (3)$$

The signal flow of calculation of RMS current and voltage values is presented in Fig. 2.

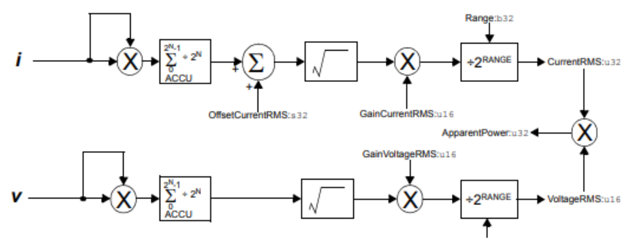


Fig. 2. RMS current and voltage calculation signal flow

The active power is obtained as average value of the multiplied instantaneous values of current and voltage, by using Eq. (4). Signal flow of active power calculation is shown in Fig. 3. The MCP39F521 has two simultaneous sampling A/D converters. For active power calculation, the instantaneous currents and voltages are multiplied in order to create instantaneous power. The instantaneous power is then converted to active power by averaging or calculating DC component.

$$P = \frac{\sum_{n=1}^{N} v(nT) i(nT)}{N} = \frac{\sum_{n=1}^{N} p(nT)}{N} . \qquad (4)$$

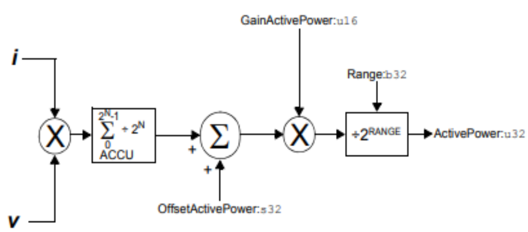


Fig. 3. Active power calculation signal flow

III. REALIZED SYSTEM

The block diagram and photography of our system are shown in Fig. 4. As can be seen from Fig. 4a our system consists relay circuit, RS485 circuit, current/voltage sensor circuit, power measuring circuit and MCU. We used a well-known microcontroller Atmega328P [10] as MCU, which characteristics meet all our demands. MCU communicates with power measuring circuit (MCP39F521) by using I2C protocol and passes the obtained data through RS485 to central monitoring system by using Modbus RTU protocol.

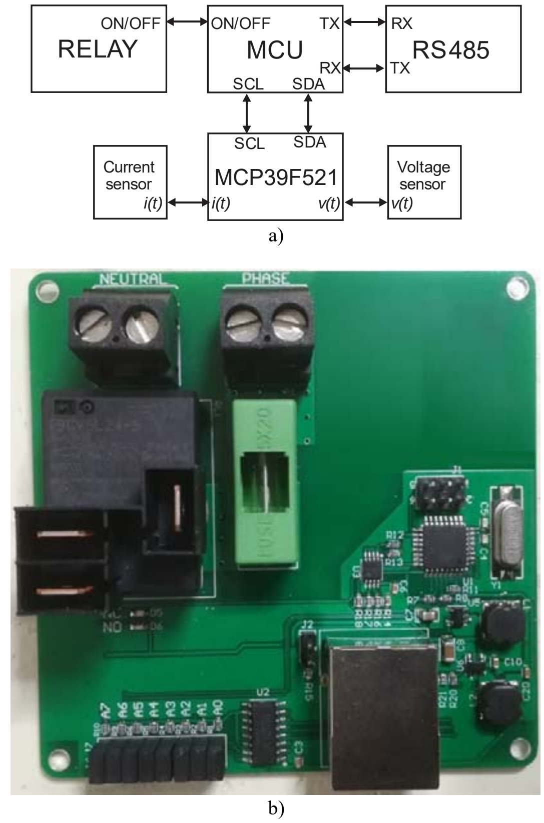


Fig. 4. a) Block diagram of realized system, b) photo of realized system

RS485 interface supports multiple devices on the same bus. This interface is currently widely used in data acquisition and control applications where multiple nodes communicate with each other. Consequently, each board needs to have unique address (unique on the network level). Because of that remote monitoring computer can send request only to specific node in a network. This address can be set by using jumpers in our system. The maximum number of nodes on the same RS-485 network in our case is 213 [11]. The number of supported devices on same networks depends on IC that is used for RS485 interface, bound rate and distance between nodes.

The load is powered via the T9A series relay, which has normal open (NO) and normal closed terminals (NC) [12]. The relay can be set to normal closed state, which is useful in applications where electrical appliance has to be always on, but on demand it can be turned off. Typical example of this usage is refrigerator in hotel rooms which needs to be always powered on. When customer exits the room, all electrical outlets and devices are disabled, except refrigerator which is connected via the relay board.

As we said before, communication between computer and our device is done by using Modbus RTU protocol and QModMaster application. We use only four functions from Modbus RTU:*write single coil, write multiple coil, read input registers and read holding registers.*

The state of relay can be controlled by using *write single coil* or *write multiple coil* function, while reading measurement results is done by using *read input registers* or *read holding registers* functions. These registers contain value of RMS voltage/current, line frequency, power factor, active power and total accumulated energy, as shown in Fig.5.

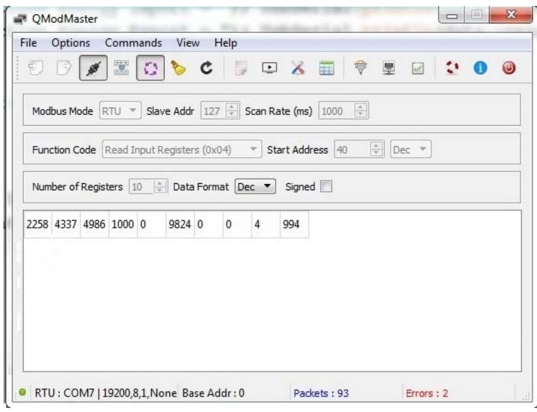


Fig. 5. Modbus application

QModMaster application runs on central computer which is used for reading data from each relay board. For connecting PC to RS-485 network USB-RS485 bridge is used. Before connecting relay board on the network, it is necessary to set slave address of the board, by setting appropriate jumper configuration. In Fig. 5, we showed obtained data when using *read input registers* function. By using this function and setting number of register to 10 we obtain: RMS voltage/current (registers 1 and 2), line frequency (register 3), power factor(register 4), active power(registers 5 and 6) and active energy(registers 7 to 10), respectively. In order to get good accuracy of read values, we introduce correction factors. This factor for voltage is 10, for line frequency and active power is 100, while for current, power factor and active consumed energy is 1000. The negative values of power factor are presented with digit 1 on the 5th position at register 4.

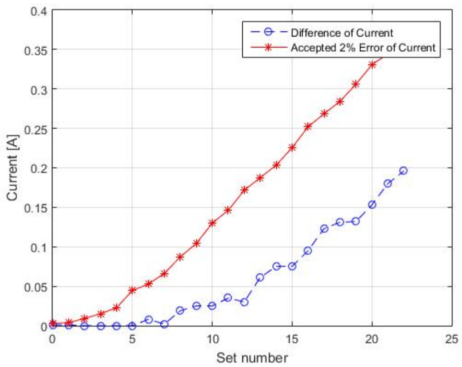
IV. Measurement results

The accuracy of the realized system is verified by using a set of different linear and nonlinear loads. As nonlinear loads we chose LED and CFL bulbs. These nonlinear loads are chosen as benchmarks because they characterize small nominal power. Namely, the intention is to show that our system measures current in range from 0.1A to 20A with accuracy less than 2%, having different load conditions. For linear load measuring we have used industrial resistor whose resistance can be set in range from 16.66 **Ω** to 500 **Ω**, heater and different types of incandescent light bulbs (ILB). As reference measuring instrument, we have used electronic power meter produced by EWG electronics [13]. It fulfils the IEC 62052-22 standard [14], IEC 62052-23standard [15]. The previously mentioned standards, fulfilled by power meter, guaranty to us that power meter has good accuracy.

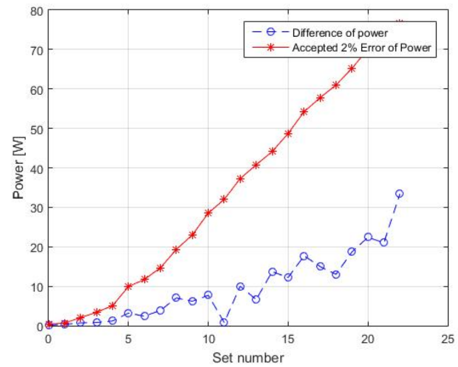
TABLE I
Measurement results

		Power Meter			**Our System**					
NO.	TYPE OF LOAD	U_{RMS}	I_{RMS}	P (W)	U_{RMS}	I_{RMS}	P (W)	V_{RMS} Error(%)	I_{RMS} Error(%)	P Error(%)
1	CFL20W	225	0.135	17.7	224.8	0.134	17.8	0.09	0.75	0.56
2	LED10W+CFL20W	224.8	0.193	32.72	224.7	0.194	32.4	0.04	0.52	0.99
3	R=500Ω	225	0.455	102	225.00	0.455	102.70	0.00	0.00	0.68
4	R=300Ω	224.7	0.758	169.9	224.70	0.758	170.70	0.00	0.00	0.47
5	R=200Ω	224.6	1.131	253	224.20	1.131	254.20	0.18	0.00	0.47
6	R=100Ω	223.5	2.23	494.6	222.90	2.230	497.80	0.27	0.00	0.64
7	R=500Ω&ILB100W	223.1	2.632	586.2	222.6	2.640	588.6	0.22	0.30	0.41
8	R=70Ω	222.6	3.309	733.33	222.1	3.311	737.2	0.23	0.06	0.52
9	R=50Ω	221.44	4.368	966.37	221.1	4.387	973.4	0.15	0.43	0.72
10	R=50Ω&ILB200W	220.09	5.207	1150.2	220.05	5.232	1156.4	0.02	0.48	0.54
11	R=35Ω	219.8	6.495	1425.6	219.1	6.520	1433.3	0.32	0.38	0.54
12	R=35Ω&ILB200W	218.8	7.311	1600.6	218.2	7.346	1601.4	0.27	0.48	0.05
13	R=25Ω	217.6	8.574	1864	217	8.604	1874	0.28	0.35	0.53
14	R=25Ω&ILB200W	217.44	9.375	2038.9	216.5	9.436	2045.5	0.43	0.65	0.32
15	R=25Ω&ILB400W	216.73	10.171	2206	215.8	10.246	2219.7	0.43	0.73	0.62
16	R=25Ω&ILB650	216.03	11.272	2435.5	215	11.347	2447.7	0.48	0.66	0.50
17	R=16.66Ω	215.38	12.6	2711	214.3	12.695	2728.7	0.50	0.75	0.65
18	R=16.66Ω&ILB200W	214.7	13.424	2888	213.9	13.547	2903	0.37	0.91	0.52
19	R=16.66Ω&ILB400W	214.51	14.215	3050	213.1	14.346	3063	0.66	0.91	0.42

20	R=16.66Ω&ILB650W	212.85	15.3	3260	211.8	15.432	3278.7	0.50	0.86	0.57
21	R=16.66Ω&Heater&ILB200W	212.56	16.533	3506	210.9	16.686	3528.4	0.79	0.92	0.63
22	R=16.66Ω&Heater&ILB400W	211.4	17.22	3641	209.3	17.400	3662	1.00	1.03	0.57
23	R=16.66Ω&Heater&ILB650W	210.7	18.25	3833	209.1	18.446	3866.4	0.77	1.06	0.86



a)



b)

Fig. 6. a)Accuracy of current for load shown in Table I , b) Accuracy of power for load shown in Table I

As it is shown in Table I and Fig 6.,after measurements with all the loads, we obtained accuracy less than 2%.

V. Conclusion

This paper presented a current and power monitoring system with accuracy less than 2%. Measuring of voltage, current, line frequency and active power for household appliances gives us sufficient information on which we can perform some action. Built-in relay with NO and NC contacts drastically increases practical usability of our system, so that remote controlling of appliance is also supported. This is very important in terms of creating a power scheme, so that some device can be powered only during night hours, or in case of some unpredicted behaviour we can switch off corresponding device, etc. In the future our goal will be to expand monitoring capabilities so that reactive and apparent power can be measured. Also, beside RS-485 interface, adding WiFi and/or Bluetooth capability will affect places where physically adding cables for RS-485 interface is not an option.

Acknowledgment

This work has been supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

References

[1] "Energy Efficiency and its contribution to energy security and the 2030 Framework for climate and energy policy", https://ec.europa.eu/energy/sites/ener/files/documents/2014_eec_communication_adopted_0.pdf

[2] M.Clemence, R.Coccioni, A.Glatigny, "How utility electrical distribution networks can save energy in the smart grid era", Schneider Electric, April 2013

[3] M.AndrejevićStošović, D.Stevanović, P.Petković, "*Application of a Standard Power Meter for Detection Source of Harmonic Pollution and Reducing Economic Losses at Power Grid*, Electric Power Components and Systems, Vol. 48, Issue 1-2, Taylor & Francis, UK, 07.04., 2020, pp. 42-55, ISSN 1532-5008, doi:10.1080/15325008.2020.1731879

[4] D.Stevanović, M. AndrejevićStošović, M. Dimitrijević "*Improving the system for registration of electric energy consumption*, 2020 International Symposium on Industrial Electronics and Applications " INDEL 2020, Banja Luka, Bosnia and Hercegovina, 04.11.-06.11., 2020, pp. 1-5, ISBN 978-1-7281-9863-7, doi:10.1109/INDEL50386.2020.9266207

[5] https://nerdtechy.com/best-smart-home-energy-monitor

[6] J. G. Josué, J. M. Pina, M. V. Neves "Home Electric Energy Monitoring System: Design and Prototyping",Proceedings of Technological Innovation for Sustainability – Second IFIP WG 5.5/SOCOLNET Doctoral Conference on Computing 2011, Costa de Caparica, Portugal, 21 – 23. February, 2011, pp. 437 - 444.

[7] E. Barakat, N. Sinno, C. Keyrouz,"*A Remote Monitoring System for Voltage,, Current, Power and Temperature Measurements*", vol 55, no.1, pp. 421 – 428, 2014.

[8] M. J. Manti,A. V Bossche, R. F. Chisab "A Smart Voltage and Current Monitoring System for Three Phase Inverters Using an Android Smartphone Application", Sensors 2017, DOI:10.3390/s17040872

[9] https://ww1.microchip.com/downloads/en/DeviceDoc/20005442A.pdf

[10] https://ww1.microchip.com/downloads/en/DeviceDoc/ATmega48A-PA-88A-PA-168A-PA-328-P-DS-DS40002061B.pdf

[11] https://www.ti.com/lit/ds/symlink/sn65hvd72.pdf?HQS=dis-dk-null-digikeymode-dsf-pf-null-wwe&ts=1621931109406

[12] https://www.te.com/commerce/DocumentDelivery/DDEController?Action=srchrtrv&DocNm=1308242_T9A&DocType=DS&DocLang=English

[13] EWG - multi metering solutions, www.ewg.rs

[14] IEC 62053-22 Electricity metering equipment (AC) - Particular requirements - Static meters for active energy (classes 0.2S and 0.5S).

[15] IEC62053-23 (2003) Electricity metering equipment (a.c.) – Particular requirements – Part 23: Static meters for reactive energy (classes 2 and 3)

Matlab/Simulink 1D model of longitudinal wave propagation through piezoceramic rings

Igor Jovanović and Dragan Mančić

***Abstract*— One-dimensional (1D) model of piezoelectric elements enable fast prediction of performance, as well as a good insight into the behaviour of piezoelectric elements and the entire ultrasonic transducer during operation. In this paper, the Matlab/Simulink 1D model of piezoceramic rings that include only thickness oscillation modes is presented, while radial oscillations are neglected. Implementation of equivalent electromechanical circuits in the modelling of piezoelectric elements did not bring a larger amount of information in relation to the number of information obtained by applying constitutive piezoelectric equations. In addition, the presented model that directly relies on the constitutive piezoelectric equations enables better visualization of wave propagation through transducer structure. The input electrical impedances for piezoceramic rings are calculated using the realized model and then compared with experimental results to validate the model.**

***Index Terms*— Piezoceramic ring, Matlab/Simulink model, Input electrical impedance.**

I. Introduction

Models involving computer simulations have become an essential part of the transducer design process. Electromechanical equivalent circuits are frequently used to model and analyse ultrasonic transducers, whose application is based on the idea that wave propagation speed is equivalent to electric current, while mechanical force is equivalent to electric voltage [1]. Today, many transducers are designed with one dominant resonant mode, which can further simplify the models and justify their use. Piezoceramic rings of different thicknesses and inner/outer diameters are widely used as active components in ultrasonic sandwich piezoceramic transducers [2].

In case that the axial dimensions of ultrasonic transducers that oscillate in the thickness mode are larger than the radial dimensions, one-dimensional analysis can be applied in the modelling process of both piezoceramic rings and whole sandwich transducers [3], [4]. Although this dimensional relationship is common with most ultrasonic transducers, in the process of modelling transducers with a complex structure (e.g. composite transducers), it is important to predict the behaviour of transducers in all directions of oscillation propagation [5]. In this case, it is necessary to use three-dimensional models [6]. The components of these power transducers can be modelled by mathematical analysis with the help of appropriate physical laws.

Igor Jovanović and Dragan Mančić are with the University of Niš, Faculty of Electronic Engineering, 14 Aleksandra Medvedeva, 18000 Niš, Serbia (e-mail: igor.jovanovic@elfak.ni.ac.rs and dragan.mancic@elfak.ni.ac.rs).

In this paper, the Matlab/Simulink model of piezoceramic rings based on constitutive equations of piezoelectric effect is presented. The proposed Matlab/Simulink model leads to simpler implementation than the mathematical model. This model can also be used to analyse multilayer structures that include both piezoelectric materials and metal endings.

II. Description of Governing Equations

Constitutive equations for piezoelectric material can be written in the following form when neglecting transverse dimensions [7]:

$$E = -hS + \frac{D}{\varepsilon^S}, \tag{1}$$

$$T = c^D S - hD, \tag{2}$$

where the mechanical stress T can be determined by dividing the total extension or compression force F by the transverse surface P, $T=F/P$. E is the applied electrical field, S is the mechanical deformation, D is the dielectric displacement, c^D is the elastic stiffness coefficient, h and ε^S describe physical characteristics of piezoelectric material.

By applying Newton's II law, which defines force as a product of mass and acceleration, and using mass as a product of volume and density, mechanical stress is expressed as:

$$T = \rho l \frac{\partial^2 u}{\partial t^2}, \tag{3}$$

where ρ is the density, l is the thickness of the piezoelectric material, and u are the mechanical displacements components. By differentiating the last equation along the z-axis, it is obtained [8]:

$$\frac{\partial T}{\partial z} = \rho \frac{\partial^2 u}{\partial t^2}. \tag{4}$$

Expressing Hook's law in the following form [9]:

$$T = c^D \frac{\partial u}{\partial z}, \tag{5}$$

where the elastic stiffness coefficient c^D represents the coefficient of proportionality. The previous expression can be rewritten using the propagation speed of longitudinal waves in a thin cylinder $v^2=c/\rho$, as:

$$\frac{\partial^2 u}{\partial t^2} = v^2 \frac{\partial^2 u}{\partial z^2}. \tag{6}$$

In 1D models, only mechanical deformation and divergence of the dielectric displacement along the z-axis are of interest:

$$S = \frac{\partial u}{\partial z}, \text{ and } \nabla D = \frac{\partial u}{\partial z} = 0. \tag{7}$$

By differentiating expression (2) along the z-axis, one may obtain:

$$\frac{\partial T}{\partial z} = c^D \frac{\partial^2 u}{\partial z^2} - h\frac{\partial D}{\partial z}. \tag{8}$$

By solving (6) in the general case, the equation of nonstationary wave motion is obtained in the form of a linear differential equation of the second-order with two independent variables (by z-axis and by time t).

If a multilayer structure shown in Fig. 1(a) is observed, the solution of this hyperbolic partial differential equation (determinant is $4v^2 > 0$) can be written in general form for layer n in the multilayer structure:

$$u_n = \alpha_n e^{-j\omega\Delta t_n} + \beta_n e^{j\omega\Delta t_n}. \tag{9}$$

Amplitudes of the mechanical displacements components in the direction of wave propagation are denoted by α_n, and in the opposite direction to the direction of wave propagation by β_n, on the surface of the n-th layer. The ultrasonic wave propagation time through the n-th piezoelectric layer is equal to the ratio of the length of the propagation path and the speed of wave propagation (for propagation through the entire layer $\Delta t_n=l_n/v_n$).

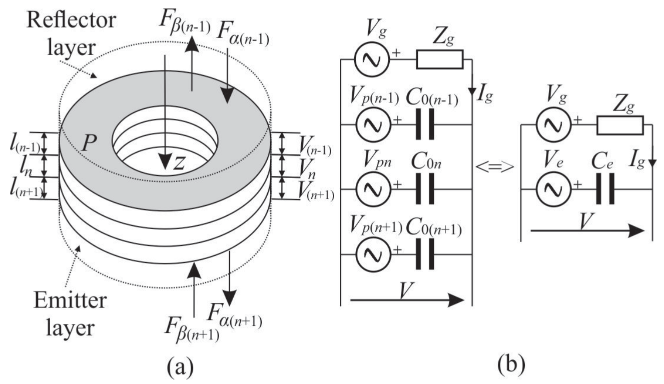


Fig. 1. (a) Multilayer structure of an ultrasonic sandwich piezoceramic transducer, (b) its equivalent circuit in the case of a parallel electrical excitation.

To obtain the analytical expression for the electric voltage V_n on the n-th piezoelectric layer, (1) can be written in the following form:

$$E_n = -h_n \frac{\partial u_n}{\partial z} + \frac{Q_n}{P\varepsilon_n^S}, \tag{10}$$

for $D=Q_n/P$ and static capacitance of the n-th layer $C_{0n}=\varepsilon_n^S P/l_n$. Q_n is the amount of accumulated charge stored on capacitor C_{0n}.

By integrating the last equation along the z-axis the following expression was obtained [9]:

$$\begin{aligned} V_n &= -h_n \int_0^{l_n} \frac{\partial u_n}{\partial z} dz + \frac{Q_n}{P\varepsilon_n^S} \int_0^{l_n} dz = \\ &= -h_n \left[\alpha_n \left(e^{-j\omega\Delta t_n} - 1\right) + \beta_n \left(e^{-j\omega\Delta t_n} - 1\right)\right] + \frac{Q_n}{C_{0n}}. \end{aligned} \tag{11}$$

Is useful to introduce mechanical displacements components u_n with the help of forces acting on the surfaces, which are perpendicular to the z-axis ($F_{\alpha n}$ and $F_{\beta n}$) [9]:

$$V_n = -\frac{h_n}{j\omega Z_{cn}} \left(F_{\alpha_n} + F_{\beta_n}\right) \left(1 - e^{-j\omega\Delta t_n}\right) + \frac{Q_n}{C_{0n}}, \tag{12}$$

where the specific acoustic impedance Z_{cn} and the amplitudes of the mechanical displacements components are represented as:

$$Z_{cn} = \frac{c_n^D P}{v_n}, \quad \alpha_n = -\frac{F_{\alpha_n}}{j\omega Z_{cn}} \text{ and } \beta_n = \frac{F_{\beta_n}}{j\omega Z_{cn}} e^{-j\omega\Delta t_n}. \tag{13}$$

The general analysis of the multilayer structure of an ultrasonic sandwich piezoceramic transducer includes several piezoelectric layers that act as sources or as sensors of oscillations. However, in this paper, the analysis is limited to the most commonly used electrical connection of piezoelectric layers in practice, to the analysis of single-layer or multilayer structures that are mechanically connected in series and electrically excited in parallel, Fig. 1(a). Fig. 1(b) shows that the voltage V of the parallel connection of all active layers consists of voltage contributions based on the piezoelectric effect V_{pn} and the amount of accumulated charge stored on capacitor C_{0n}. The equivalent capacitance of the parallel connection of N piezoelectric layers C_e, as well as the total amount of charge Q_e, which depends on the voltage of the electric generator V_g, and its impedance Z_g, can be represented by the expressions:

$$C_e = \sum_{n=1}^{N} C_{0n}, \quad Q_e = \frac{V_g - V}{j\omega Z_g}. \tag{14}$$

When the equivalent circuit of Fig. 1(b) is considered as a capacitive voltage divider, the equivalent voltage generated by the piezoelectric effect can be calculated as follows [10]:

$$V_e = \frac{V_g}{1 + j\omega Z_g C_e} - \\ - \frac{j\omega Z_g C_e}{1 + j\omega Z_g C_e} \sum_{n=1}^{N} \frac{C_{0n}}{C_e} \frac{h_n}{j\omega Z_{cn}} \left(F_{\alpha_n} + F_{\beta_n} \right) \left(1 - e^{-j\omega \Delta t_n} \right). \quad (15)$$

If the structure has only one active layer, then the voltage on it is calculated by the expression:

$$V = \frac{1}{1 + j\omega Z_g C_0} \left[V_g - j\omega Z_g C_0 \frac{h_n}{j\omega Z_c} \left(F_\alpha + F_\beta \right) \left(1 - e^{-j\omega \Delta t} \right) \right]. \quad (16)$$

It is necessary to present mechanical deformation S in (2) as a derivative of mechanical displacements components u along the z-axis (7), and use expressions (13) for amplitudes of the mechanical displacements components to obtain analytical expressions for forces acting on piezoceramic surfaces [11]:

$$F_n = c_n^D P \frac{j\omega}{v_n} \left(\frac{F_{\alpha_n}}{j\omega Z_{cn}} e^{-j\omega \Delta t_n} + \frac{F_{\beta_n}}{j\omega Z_{cn}} \right) - h_n Q_n. \quad (17)$$

Equation (17) is an expression for the force on the external surfaces (perpendicular to the z-axis) of the n-th layer in the observed piezoceramic structure, in general form. Boundary conditions can be expressed based on continuity of forces and mechanical displacements components on contact surfaces:

$$F_n \big|_{z=l_n} = F_{n+1} \big|_{z=0} \text{ and } u_n \big|_{z=l_n} = u_{n+1} \big|_{z=0}. \quad (18)$$

When the ultrasonic wave passes from one layer to another, the speed of propagation changes. For explicit analysis of forces acting on the boundary surfaces between two layers, it is necessary to define the transmission and reflection coefficients during the passage of an ultrasonic wave from one layer (n) to another (n+1) [12]:

$$T_n^{n+1} = \frac{2 Z_{c(n+1)}}{Z_{c(n+1)} + Z_{cn}} \text{ and } R_n^{n+1} = \frac{Z_{c(n+1)} - Z_{cn}}{Z_{c(n+1)} + Z_{cn}}. \quad (19)$$

Components of mechanical forces acting on boundary surfaces, with defined boundary conditions and taking into account the transmission and reflection of the ultrasonic waves, can be presented by the following expressions:

$$F_{\alpha_n} = \frac{1}{1 - T_{n-1}^n K_n} \Big[F_{\alpha_{(n-1)}} T_{n-1}^n \left(e^{-j\omega \Delta t_{(n-1)}} - K_{n-1} \right) - F_{\beta_{(n-1)}} T_{n-1}^n K_{n-1} + \\ + F_{\beta_n} \left(R_n^{n-1} e^{-j\omega \Delta t_n} + T_{n-1}^n K_n \right) + \\ + \frac{T_{n-1}^n V}{2} \left(h_n C_{0n} - h_{n-1} C_{0(n-1)} \right) \Big], \quad (20)$$

$$F_{\beta_n} = \frac{1}{1 - T_{n+1}^n K_n} \Big[F_{\beta_{(n+1)}} T_{n+1}^n \left(e^{-j\omega \Delta t_{(n+1)}} - K_{n+1} \right) - F_{\alpha_{(n+1)}} T_{n+1}^n K_{n+1} + \\ + F_{\alpha_n} \left(R_n^{n+1} e^{-j\omega \Delta t_n} + T_{n+1}^n K_n \right) + \\ + \frac{T_{n+1}^n V}{2} \left(h_n C_{0n} - h_{n+1} C_{0(n+1)} \right) \Big], \quad (21)$$

wherein $K_n = h_n^2 C_{0n} \frac{1 - e^{-j\omega \Delta t_n}}{2 j\omega Z_{cn}}$.

Indices in (20) and (21) indicate that for the first layer (n = 1) and the last layer (n = N), the values with indices $n - 1 = 0$ and $n + 1 = N + 1$, refer to the propagation medium behind the reflector layer and in front of the emitter layer, respectively.

The components of mechanical forces on the surfaces of the unloaded piezoceramic transducer that consists of only one piezoceramic layer (the influence of external forces on the transducer is equal to zero $F_{\alpha 0}$=$F_{\beta 2}$=0) are calculated as:

$$F_{\alpha_1} = \frac{1}{1 - T_0^1 K} \left[F_{\beta_1} \left(R_1^0 e^{-j\omega \Delta t_1} + T_0^1 K \right) + \frac{T_0^1 V}{2} h C_0 \right], \quad (22)$$

$$F_{\beta_1} = \frac{1}{1 - T_2^1 K} \left[F_{\alpha_1} \left(R_1^2 e^{-j\omega \Delta t_1} + T_2^1 K \right) + \frac{T_2^1 V}{2} h C_0 \right]. \quad (23)$$

Since this paper shows the modelling of only one active layer, which is unloaded (surrounding medium is air), external acoustic impedances are 400 Rayl, [11]. The value used for external acoustic impedances is much less than the specific acoustic impedance of the active layer, so it can be considered T_n^{n+1} = 2 and R_n^{n+1} = -1.

III. Simulation and Experimental Results

Expressions (15), (20) and (21) can form a system of equations that describes the electromechanical structure shown in Fig. 1. By solving this system of equations, numerical results are obtained that represent the mechanical forces in each of the layers of the packet transducer. In addition to knowing the characteristics of the electric generator, it is possible to calculate the input electrical impedance of the transducer itself.

Fig. 2 shows the Matlab/Simulink 1D model of one piezoceramic layer (a model of longitudinal wave propagation through piezoceramic rings). The model represents a system of equations formed by the Laplace transform (with related expressions for impulse delays $e^{-\tau s}$, differentiations s, and integrations $1/s$ with respect to time) of terms (16), (22) and (23).

The calculated and experimental results are obtained using a piezoceramic equivalent material. The dimensions of the used piezoceramic rings are given in the Table I, where l is the thickness, b and a are the inner and outer diameters of lead zirconate titanate (PZT) piezoceramic rings.

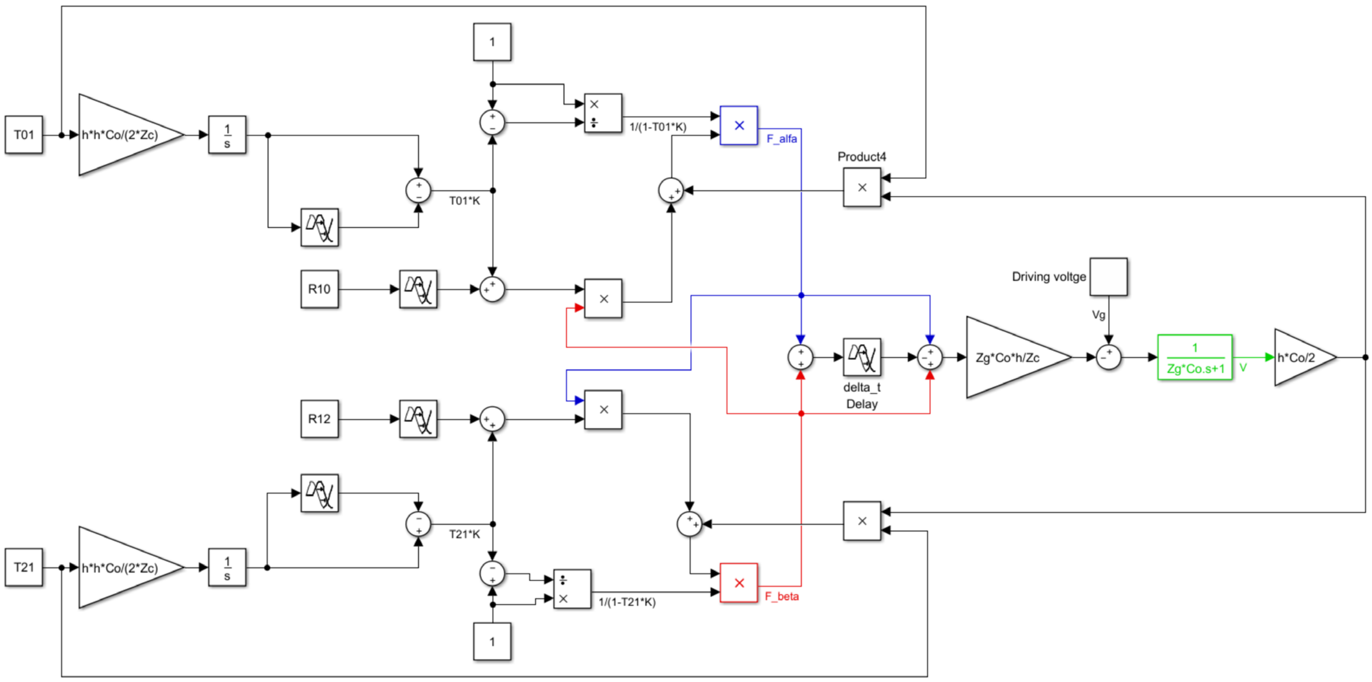


Fig. 2. Matlab/Simulink model of the piezoceramic ring.

Three samples of commercial PZT4 rings and three samples of commercial PZT8 rings have been characterized [13]. The electrical impedance measurements are conducted using an HP 4194A Network Impedance Analyzer.

TABLE I
PIEZOCERAMIC RING DIMENSIONS

Sample	*a* (mm)	*b* (mm)	*l* (mm)	*PZT equivalent material* [11]
I	38	15	5	PZT4
II	38	13	6.35	PZT4
III	50	20	6.35	PZT4
IV	24	15	3	PZT8
V	38	13	6	PZT8
VI	10	4	2	PZT8

As shown in Figs. 3-8, the measured frequency characteristic corresponds with the simulated curves using the proposed model. The proposed 1D model predicts only thickness modes while does not consider other mods. The images show the first thickness resonant mode for all used PZT samples.

The model acts as a three-port network whose ports refer to mechanical forces (F_{alfa} and F_{beta}) acting on the circular-ring surfaces of the piezoceramic ring, and the third port represents the electrical driving voltage (V_g). The charge components are proportional to the difference of mechanical displacements components between circular-ring surfaces and can be obtained including the equation (14) in the model.

These charge components further cause secondary forces to propagate through the transducer and the surrounding medium. The secondary piezoelectric effect was modelled using two positive feedback loops in Fig. 2.

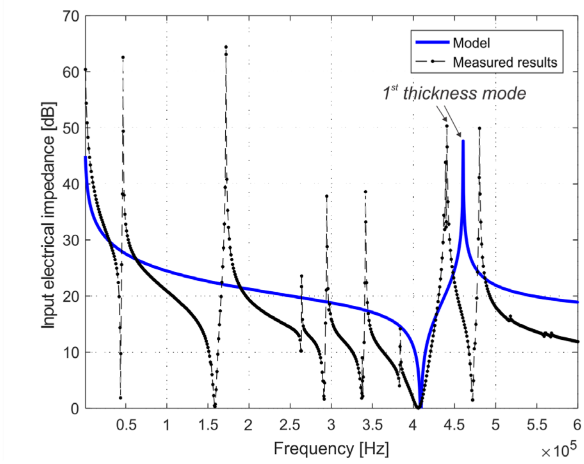


Fig. 3. Simulated and experimental input electrical impedance versus frequency for the I sample.

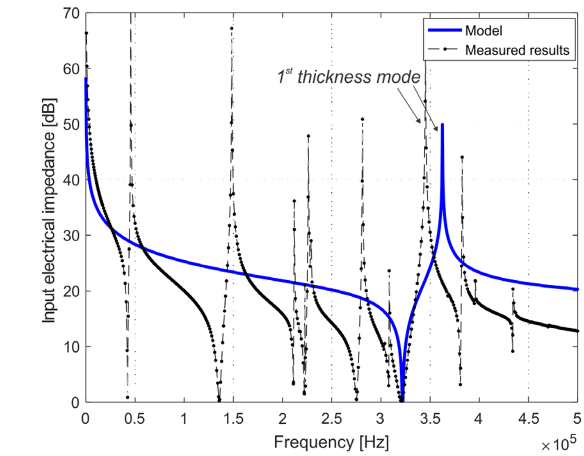


Fig. 4. Simulated and experimental input electrical impedance versus frequency for the II sample.

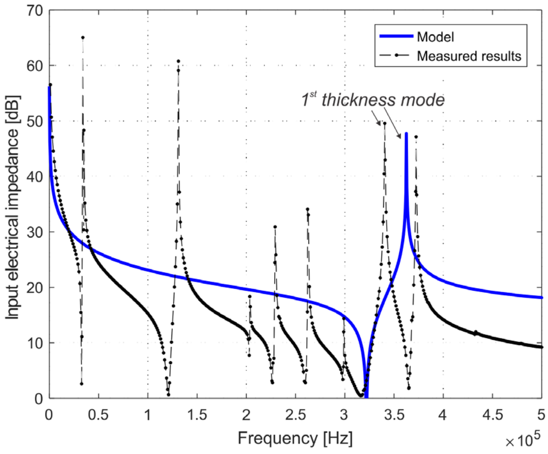


Fig. 5. Simulated and experimental input electrical impedance versus frequency for the III sample.

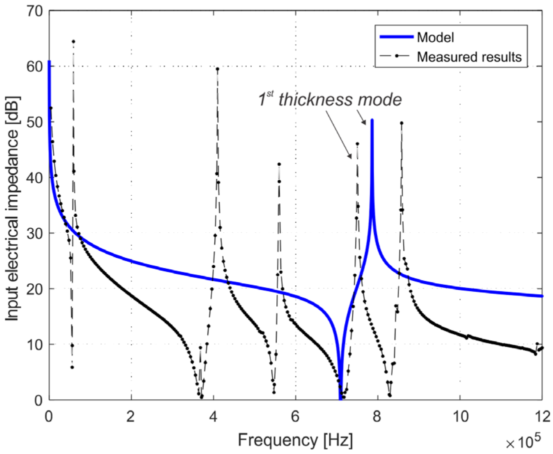


Fig. 6. Simulated and experimental input electrical impedance versus frequency for the IV sample.

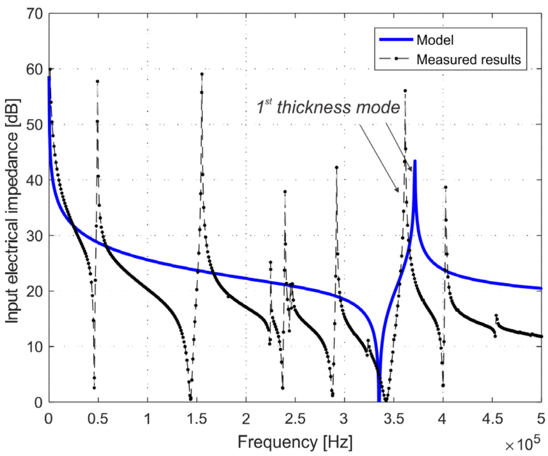


Fig. 7. Simulated and experimental input electrical impedance versus frequency for the V sample.

The driving current I_g from Fig. 1(a) was obtained by differentiating the charge components $I_g=sQ$. The input electrical impedance was obtained by dividing the voltage V from expression (16) and driving current.

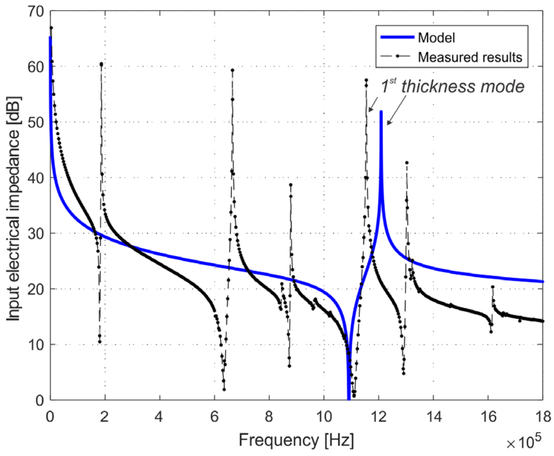


Fig. 8. Simulated and experimental input electrical impedance versus frequency for the VI sample.

IV. Conclusion

The piezoceramic rings with different dimensions are analysed using the developed model. Verification of the proposed model is performed by comparing the modelled dependencies of input electrical impedance vs. frequency with the experimental results. The matching of experimental and theoretical results is quite good and validates the proposed model. The presented model predicts thickness oscillatory modes of PZT piezoceramic rings taking the interaction with the surrounding media into account. The model describes the behaviour of the piezoceramic ring with two mechanical ports (one for each external surface normal to the z-axis) and one electrical port.

This approach is suitable for the analysis of the completely multilayer structure of the transducer. When calculating the forces acting on the passive layers of the transducer (materials that do not have piezoelectric properties), it is necessary to adopt that h=0.

This approach is not only effective in terms of computation time but also in reducing difficulties associated with the calibration of material parameters. Errors in predicting thickness oscillatory modes can be reduced by fitting the parameters of the piezoceramic material.

Acknowledgment

This work has been supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

References

[1] D. Mančić, M. Radmanović, “Piezoceramic ring loaded on each face: a three-dimensional approach," Electronic Journal "Technical acoustics", vol. 2, pp. 1.1-1.7, 2002.

[2] I. Jovanović, D. Mančić, V. Paunović, M. Radmanović, Z. Petrušić, “A Matlab/Simulink Model of Piezoceramic Ring for Transducer Design," ICEST, Niš, Serbia, vol. 3, pp. 952-955, 2011.

[3] W. P. Mason, *Electromechanical Transducers and Wave Filters*, Princeton: Van Nostrand, 1948.

[4] A. Abdullah, M. Shahini, A. Pak, “An approach to design a high power piezoelectric ultrasonic transducer," J. Electroceram. vol. 22, pp. 369–382, 2009.

[5] I. Jovanović, D. Mančić, U. Jovanović, M. Prokić, “A 3D model of new composite ultrasonic transducer, "Journal of Computational Electronics, vol. 16, no. 3, pp. 977-986, 2017.

[6] D. Mančić, G. Stančić, “New Three-dimensional Matrix Models of the Ultrasonic Sandwich Transducers, "Journal of Sandwich Structures & Materials, vol. 12, pp. 63-80, 2010.

[7] Z. H. Zhang, B. Y. Sun, L. P. Shi, “Theoretical and Experimental Study on Secondary Piezoelectric Effect Based on PZT-5,” Phys.: Conf. Ser. vol. 48, pp. 86-90, 2016

[8] U. Jovanović, Lj. Perić, I. Jovanović, D. Mančić, “Analisys of longitudinal oscillations of Free prismatic piezoceramic beams", 24th International Conference "Noise and Vibration", Niš, Serbia, pp. 229 - 235, 29-31. October 2014.

[9] S. Cochran, C. E. M. Demore, C. R. P. Courtney, “Modelling ultrasonic transducer performance: one dimensional models," in *Ultrasonic transducers: materials and design for sensors, actuators and medical applications*, Cambridge: Woodhead Publishing Ltd, 2012. pp. 187-219.

[10] J. L. Dion, E. Cornieles, F. Galindo, K. Agbossou, “Exact one-dimensional computation of ultrasonic transducers with several piezoelectric elements and passive layers using the transmission line analogy, "IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control, vol. 44, no. 5, pp. 112031, 1997.

[11] I. Jovanović, “New Power Composite Ultrasonic Transducers,” Ph.D. dissertation (in Serbian), University of Niš, Faculty of Electronic Engineering, Niš, Serbia, 2018

[12] G. Hayward, M. N. Jackson, “A lattice model of the thickness-mode piezoelectric transducer," IEEE Trans. Ultrason. Ferroelec. Freq. Contr. Vol. 33, No. 1, pp. 41-50, 1986.

[13] *Technical Publication TP-226*, Morgan Electro Ceramics.

Arduino-Based Gas and Smoke Detector Realized Using MQ-2 Sensor

Petar Stančić, Aleksandra Stojković and Miljana Milić *Member, IEEE*

***Abstract*— System for detecting carbon monoxide, particles of smoke and combustible gases is described in this paper. Detector is based on application of Arduino platform with MQ-2 sensor as a source of input signal. The gas sensor functioning principles and detector configuration are explained in details. The described detector could be successfully used for various gas leakage detection, alone, or as a part of more complex system. The experimental results are given for burning paper smoke and three types of combustible gases. They confirm good performance of the system.**

***Index Terms*—Arduino, microcontroller, flammable gas detector.**

I. Introduction

We are the witnesses of rapid growth and wide application of electronic devices and systems in everyday life. The new disciplines and branches of technical sciences have emerged, which are now developing at an even faster pace. Electronic, mechanical and computer science engineering offer implementation of components that enable the development of complex systems that have a wide range of applications in industry, medicine and other fields of human life and work. Protection of human lives and health becomes a primary goal in modern conditions. The Arduino platform has an enormous potential as an educational and research tool, and represents an excellent base for designing important real-life systems. By using the Arduino platform, which consists of a physical part with a microcontroller and of a software, it is possible to design a system that is capable of protecting a human life and goods from fire and combustible gases.

II. System components

Key components for gas/smoke detector system are J1 piezo buzzer, a green and a red LED diode, an MQ-2 smoke detector and Arduino platform. Some of them will be explained in more details next.

A. Arduino

Arduino is an open-source electronics platform, based on easy-to-use hardware and software [1]. Over the years, it has been used for thousands of projects, from everyday circuits to complex scientific instruments. The entire Arduino project has started in 2004. when a Colombian student made a “Wiring” platform for his graduate thesis. In this way a new, low cost, and simple electronic device for fast prototyping was created. Arduino programs are written using a simplified version of C++, which makes it easier to learn. Arduino boards are very versatile and can be used for a variety of different applications. Some of them are: Uno, Due, Mega, Leonardo, Micro, Esplora etc. For the purpose of gas detection system, Arduino Uno was considered quite acceptable. Arduino Uno is the most frequently used variant, since it is very beginner friendly. It consists of 14-digital I/O pins, where 6-pins could potentially be used for the Pulse Width Modulation (PWM) outputs, 6-analog inputs, a reset button, a power jack, a USB connection, In-Circuit Serial Programming (ICSP) header etc., and – ATMega328 [2]. ATMega328 is a high performance AVR microcontroller with 8-bit RISC (*Reduced Instruction Set Computer)* architecture. It has low power consumption and can execute 131 instructions per single clock cycle. It has 32KB ISP (*In-System Programming*) flash memory with read-while-write capabilities, 2KB SRAM, 1KB EEPROM and maximum operating frequency of 20MHz.

B. MQ-2 gas and smoke detector

From 2013. to 2017., there were more than 350 thousand fires per year occurring only in homes, and that number is only a quarter of total number of fires. Smoke detectors are very much needed, since they can help in reducing the number of fires or at least decrease the damage done. Having any smoke detector is better than having none.

Best smoke detectors can detect smoke particles, flames and carbon monoxide. Smart smoke detectors represent a cutting-edge technology for fire safety, since they can communicate through the apps and deliver alerts to a phone or some other device or system. Smoke detectors should always have a backup power source for the case of power loss.

There are two basic types of passive smoke detectors: photoelectric and ionization [3]. Combination of these makes a dual sensor smoke alarm, which is recommended for maximum protection from both fast flaming and slower fires. Photoelectric alarms use light to detect smoke. They sense sudden scattering of light when smoke enters into the detectors chamber, which further triggers the alarm. This method of detection can detect fires that begin with long duration of smoldering aptly. Photoelectric smoke detectors respond from 15 to 50 minutes faster than ionization alarms in early stages of fire. Ionization alarms use radiation to detect smoke [4]. They carry a small amount of radioactive material

Petar Stančić, Aleksandra Stojković and Miljana Milić are with the University of Niš, Faculty of Electric Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: petar.stancic@elfak.rs, aleksmdv@elfak.rs, miljana.milic@elfak.ni.ac.rs

between two electrically charged plates, which ionizes the air and causes current to flow between the plates. The smoke disrupts the flow of ions and reduces the flow of current, which triggers the alarm. Ionization alarms are highly sensitive and respond from 30 to 90 seconds faster than photoelectric alarms to fast flaming fires. This type of detectors is more suited to rapid flaming fire outbursts, unlike the photoelectric detectors, which responds better to smoldering stages. Ionization detectors might perform better where there is risk of fast flaming fire, whereas photoelectric detectors react better to cases of slow smoldering, like electrical or furnishing fire. Dual smoke detectors include both photoelectric and ionization sensors, making them safest smoke and fire detection devices.

MQ-2 is a Metal Oxide Semiconductor (MOS) type of gas sensor, also known as chemiresistor. Chemiresistor is a material which changes its electrical resistance as a response to changes in the nearby environment. Sensors made from metal oxides require high temperatures (200 °C or higher) to operate because, in order for the resistivity to change, an activation energy must be overcome, thus leading us to the conclusion of MQ-2 being more efficient where there is risk of fast flaming fire, as the high temperature will be reached in shorter amount of time[7]. In our case, MQ-2 sensor's resistance changes when smoke or flammable gases are present. It requires 5V DC power supply and draws about 800mW of power. Beside the smoke, it can also detect Liquefied Petroleum Gas (LPG), alcohol, butane, propane, hydrogen, methane and carbon monoxide concentrations from 200 to 10 thousand ppm [5]. The sensitivity curve of MQ-2 sensor for different gases is shown in Fig. 1.

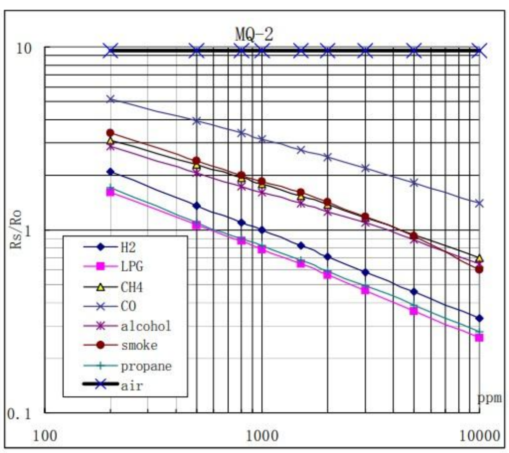


Fig. 1. MQ-2 sensitivity characteristic curve. [8]

Where R_0 is sensor resistance at 1000ppm (*parts per million*) of H_2 in the air, and R_S is sensor resistance at various concentrations of gases.

The sensor is sealed between two layers of stainless steel called anti-explosion network. It is necessary to ensure that the heater inside the sensor does not explode when flammable gases are sensed. The sensor also filters the air particles and allows that only gaseous elements pass inside the chamber. The anti-explosion network is attached to the enclosure with a copper plated clamping ring.

Fig. 2. MQ-2 external structure.

Figures 2 and 3 show the external and internal structure of the MQ-2 sensor.

Internal star-shaped structure is formed from sensing element and six connecting legs. Two of six leads, (H in Fig. 3.) are dedicated for heating the sensing element and connected through the nickel-chromium coil (conductive alloy). The remaining four leads (As and Bs in Fig. 3) generate the output signal and are connected using platinum wires which are linked to the sensing element and deliver small changes in the current that passes through the sensing element.

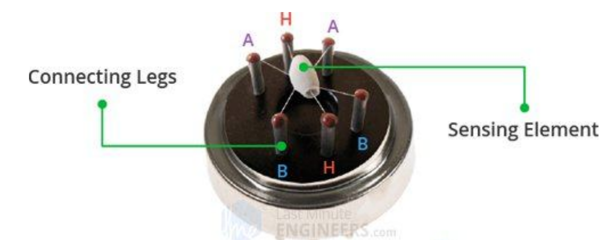


Fig. 2. MQ-2 internal structure. [9]

Sensing element is made of aluminum oxide and a tin dioxide coating. Tin dioxide coating is the most important part of the component since it is sensitive to combustible gases. Aluminum oxide increases the heating efficiency and ensures that the sensor reaches the working temperature.

If the air is clean, donor electrons from tin dioxide are attracted toward oxygen adsorbed on the surface of the sensing material, which prevents electric current flow as shown in Fig. 3. In the presence of smoke or combustible gases, oxygen reacts with gases which causes decrease of the surface oxygen density. Electrons from Fig. 5. are released back into the tin dioxide, allowing the current to flow through the sensor.

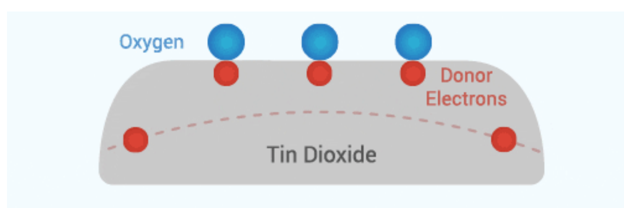


Fig. 3. Absorbed oxygen in clean air prevents the current flow.[9]

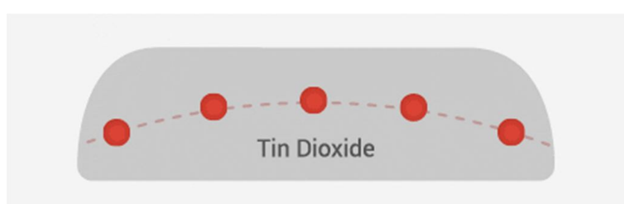


Fig. 4. Electrons are released into the tin oxide in the presence of

gas/smoke.[9]

The generated output voltage of the sensor is proportional to the concentration of the present smoke or gas. Higher gas concentration causes higher output voltage. The analog signal from MQ-2 is digitized as in the case of Arduino board based gas detection system implementation.

III. System Realization

For simulations and realization of this project, software environment "Fritzing" was used [6]. Fritzing is an open-source hardware initiative mainly used for documenting and sharing prototypes, layouts and manufacturing of Printed Circuit Boards (PCBs).

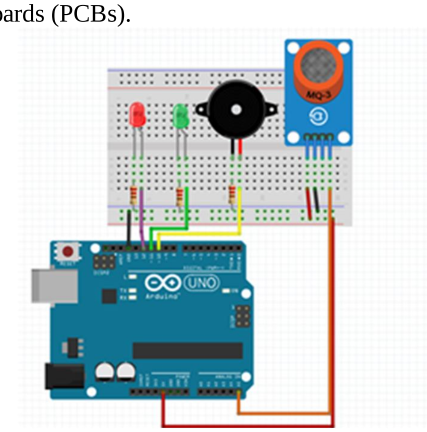


Fig. 5. Breadboard realization of the system using Fritzing.

Breadboard realization of the system using Fritzing is shown in Fig. 5.

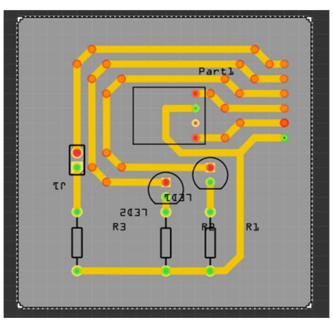


Fig. 6. PCB of our project.

PCB of the smoke/gas detection system is shown in Fig. 6.

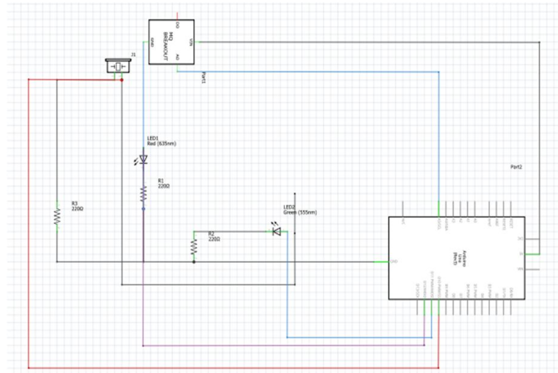


Fig. 7. Schematic of the system.

Schematic diagram of the system is shown in Fig. 7.

IV. The Experimental Results

The verification of the described detector system is performed in modest improvised laboratory environment. The greatest attention is paid to the fact that all measurements are performed in similar conditions as would be expected in real-live scenarios. In order not to endanger household safety, the measurements were performed inside a pan with a lid, as shown in Fig. 8. In this way a sufficient concentration of gases and good sensor response to small amount of gas is achieved.

Fig. 8. Experimental conditions.

Experiments show that smoke sensitivity of MQ-2 is considerably lower than its sensitivity to combustion substances (butane, alcohol and LPG in our case). This can be noticed from Fig. 10. After conducting multiple measurements, we could calculate the detector's sensitivity value (slope of the given characteristics) for different gases using the following equation:

$$u_{A5}(nT) = 5\frac{A5}{1024}\,[V], \quad T = 102\,ms \qquad (1)$$

where u_{A5} is voltage on analog pin A_5 with maximal possible value of 5V. A_5 is a digital representative of measured voltage on the sensor with maximal value of 1024.

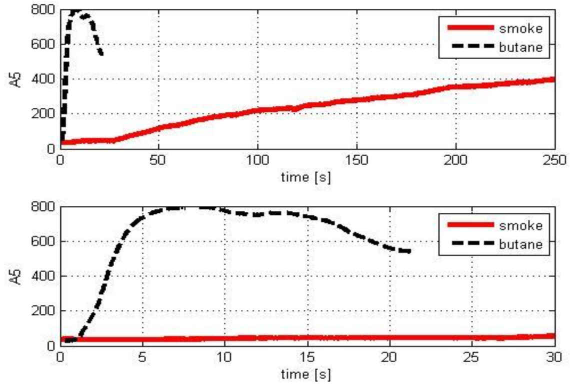


Fig. 9. Measured voltage for smoke and butane.

The relevant measure of sensor sensitivity could be the sensor's output voltage change rate $\frac{du_{A5}}{dt}$.

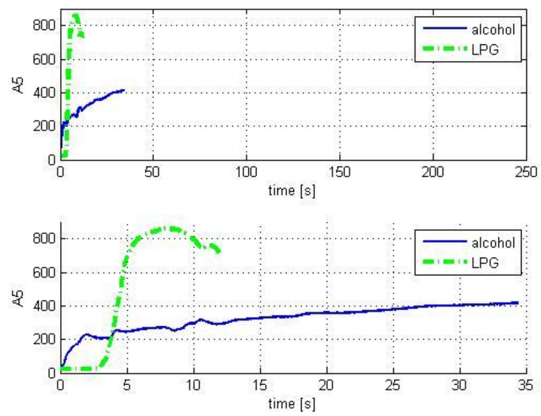


Fig. 10. Measured voltage results for alcohol and LPG.

That value is estimated from collected data, shown in Figs. 9 and 10, taking into account only part of the curve where the slope was constant and maximal.

- Butane: 0.9896 V*s^{-1}
- Alcohol: 0.5127 V*s^{-1}
- LPG: 1.8815 V*s^{-1}
- Smoke: 0.0145 V*s^{-1}

Measured voltage values were shown in *Serial Monitor* view in Fritzing, and it allowed us to track results in real time via serial data transfer (In this case it was set to 9600 bauds). Gathered data was imported in MATLAB after which it was used to generate the above characteristics. Those graphs show cumulative data. The latter graphs are zoomed and show curve parts where MQ-2 shows significant change when detecting larger amounts of gas.

One can conclude that MQ-2 is very sensitive to gases mentioned above. Considering the conditions in which experiment was conducted, obtained sensitivities cannot be considered as very accurate. The source of LPG was a 20l gas bottle. In short time it can deliver larger amounts of gas than a lighter (used as a source for butane). It is also noticed that the sensitivity to alcohol vapors is slightly lower.

Measured results are in compliance with data from Fig. 1, while maximal sensitivity is obtained for the LPG. Home conditions were not very satisfactory because the main priority was not to burn the sensor (small distance from the fire), while maintaining to "feed" the fire by removing the lid regularly. Real time information from Serial Monitor detected the smoke concentration drop with lid removal (to obtain more paper fuel) and this caused slow increase of sensor output voltage.

At the end we can conclude that with the right calibration, the MQ-2 can be used for detecting (and alarming) larger quantities of combustible gases in houses or storages.

V. Conclusion

In this paper one possible solution for realization of detector system for recognizing presence of smoke or combustible gases is presented. System is tested in home conditions with exposing MQ-2 sensor to different types of gases. Output sensor voltage is monitored for smoke generated by burning papers and cigarets and for available combustible gases as butane from lighter, stove LPG and alcohol vapors. Experiments prove high sensitivity of MQ-2 sensor making it a good choice in detector system basic sensor selection. By combining with other types of sensors, it is possible to create more complex detectors.

Acknowledgment

This work has been supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

References

[1] M. Verde, "*Pic Mikrokontroleri*", Beograd : Mikroelektronika; 2008
[2] http://www.atmel.com/devices/ATMEGA328.aspx
[3] https://www.ifsecglobal.com/smoke-detectors/
[4] D. Malain, and P. Kanchana. "Evaluation of radiation safety for ionization chamber smoke detectors containing Am-241, "*Journal of Physics: Conference Series*", Vol. 1285. No. 1. IOP Publishing, 2019.
[5] https://lastminuteengineers.com/mq2-gas-senser-arduino-tutorial/?fbclid=IwAR1gYvlT6VbUM_3bxXKsW8CnFyzmtqPiCwA6myb9DaOV_6P4SWNYBfN3sQM
[6] https://fritzing.org/
[7] O. Asif, Md. B. Hossain, M. Hasan, M. T. Rahman and M. E. H. Chowdhury, "Fire-Detectors Review and Design of an Automated, Quick Responsive Fire-Alarm System Based on SMS", *Int. J. Communications, Network and System Sciences*, Vol. 7, pp. 386-395, 2014
[8] https://www.pololu.com/file/0J309/MQ2.pdf?fbclid=IwAR3zv-g1QSDJtT4FSUAfIC3BxyQVeRiqy7dfRQjnLVBjb440svr5GfqUwWI
[9] https://lastminuteengineers.com/mq2-gas-senser-arduino-tutorial/

A Chisel Generator of JTAG to Memory-Mapped Bus Master Bridge for Agile Slave Peripherals Configuration, Testing and Validation

Vukan D. Damnjanović, *Student Member, IEEE*, and Vladimir M. Milovanović, *Senior Member, IEEE*

***Abstract*—This paper presents a design of a JTAG to memory-mapped bus master bridge generator implemented using Chisel hardware design language. This type of digital module can provide convenient and practical means of configuring a wide range of peripheral circuitry with a memory-mapped slave interface attached to a bus interconnection, as well as of their testing and debugging. The peripherals can be configured by driving the input signals of the JTAG interface with the values that represent the previously defined instruction codes, thus initiating write or read data transactions on the interconnect bus through the master interface to their memory-mapped registers. The master interface can be either AXI4 or TileLink, depending on the characteristics of the whole system which the depicted bridge is a part of. The proposed generator offers the ability of creating slightly different modules by using different parameter selections. The implemented design has been extensively tested using various software simulations with a number of different slave peripherals and mapped and tested onto a commercial FPGA platform. These actions experimentally confirmed the previously made assumption of the utility and convenience of the proposed generator.**

***Index Terms*—JTAG to memory-mapped bus master bridge, AXI4 and TileLink protocols, memory-mapped interface, peripherals testing and debugging, Chisel hardware design language, design generator.**

I. Introduction

From the very beginnings of the integrated circuits and the emergence of the first microprocessors, memories and data storing were emphasized as one of the most important and vital parts of its structure because of a vast number of possibilities and numerous functions it provided. Almost all digital applications and devices were able to develop and operate on its basis. As the time passed by, with the improvement of the existing technologies and the emergence of the new ones, along with the development of microprocessors, those applications and devices started to become more and more complex and sophisticated as well. Therefore, not only that the limited capacities of the devices' memories appeared to be the major problem, but the ways of accessing their data were too, usually due to a need for the standardized methods or high performance criteria of the systems. Several ways of a microprocessor data access were developed over the years, with the usage of the port-mapped input/output (PMIO) and the memory-mapped input/output (MMIO) interfaces [1] being the most common ones.

The main characteristic of the port-mapped input/output data access interface is the presence of special address space outside the common system memory for every included peripheral. Usually, that implied the existence of special, dedicated instruction set for data access, such as ”IN” and ”OUT” instructions in x86 architectures [2]. The PMIO was more extensively utilized in earlier digital systems with less developed microprocessors with small address spaces, since the valuable resources were not consumed by the input/output (IO) devices. However, sometimes it is not convenient to use this kind of data access because of the possible frequent context switching or the need for the manipulation of IO devices using only standardized memory access instructions. Those features are delivered by using the memory-mapped input/output interface.

As mentioned before, systems with the MMIO have a shared virtual address space, along with the program memory or user memory, with the same instruction set for accessing it. All the devices are attached to an interconnect bus and from the perspective of the microprocessor, there is no difference whether it manipulates with the peripheral I/O device, or some internal data. Therefore, a wide range of different peripherals with the memory-mapped registers can be integrated in the system without almost any additional logic, thus allowing it to grow plentifully in terms of its functionality. Having this in mind, it is no wonder that modern-day systems more and more rely on this version of peripheral device's data access.

In the previous couple of paragraphs, the characteristics and the importance of the MMIO interface within the microprocessor-based systems were elaborated. For each one of those systems, manipulating the peripheral devices by accessing their data should be well-explained and straightforward. However, sometimes there is a need to manipulate or test those devices without implicating the microprocessor. In those cases, accessing the interconnect bus and initiating data transactions could be very challenging and complex. For that particular reason, the JTAG to memory-mapped bus master bridge generator from this paper's topic was designed and created. It allows a user to access the peripheral device connected to either AXI4 [3] or TileLink [4] bus without the engagement of any kind of processing unit, but by using the standardized and quite popular JTAG interface [5], [6].

Vukan D. Damnjanović is with NOVELIC d.o.o., Veljka Dugoševića 54/B5, 11060 Belgrade, Serbia (e-mail: vukan.damnjanovic@novelic.com).

Vladimir M. Milovanović is with the Department of Electrical Engineering, Faculty of Engineering, University of Kragujevac, Sestre Janjić 6, 34000 Kragujevac, Serbia (e-mail: vlada@kg.ac.rs).

This paper, along with the quick overview of the used protocols and detailed description of the design of the JTAG to memory-mapped master bus bridge and its implementation with the used Chisel libraries, also serves as the user manual of the module and depicts the obtained results through simulations and hardware implementation.

II. A JTAG to Memory-Mapped Bus Master Bridge, Its Interfaces and Instructions

A JTAG to memory-mapped bus master bridge [7], [8] is a digital component which initiates data transactions with the peripheral devices possessing the memory-mapped input/output interface, attached to an interconnect bus. This module communicates with the outer world through two ends and three possible different interfaces. The user sends the desired instruction by driving the signals on the user-end JTAG interface, present in all the variants of the module. The instruction, if performed correctly, then initiates the appropriate transaction on the bus-end interface, which can be either AXI4 or TileLink, depending on the user's preferences or the system requirements. In the next few paragraphs, a brief overview of these interfaces is provided.

A. JTAG Interface

The Joint Test Action Group (JTAG) is a standardized four-wire serial protocol usually used for testing and debugging integrated circuits through a JTAG port. It consists of three input signals and an output signal: Test Clock (TCK), which is used as a clock signal for a JTAG controller and is independent of system clock signal; Test Mode Select (TMS), an input signal which serves as a control signal for a JTAG finite state machine (FSM), which will be discussed later; Test Data Input (TDI), an input signal that represents the input serial data for JTAG instruction and data registers; Test Data Output (TDO), the serial output data for JTAG instruction and data registers (an additional output signal exists to express the validity of the output data on the TDO pin). The on-chip JTAG Test Access Port (TAP) implements the mentioned FSM, which is used for the realization of the JTAG subpart of the module.

The JTAG FSM is used to correctly accept the JTAG interface signals and to recognize whether the arrived data values are the instruction values or the data values itself. It is driven by the rising edge of the TCK signal and the state changing is controlled by the TMS signal. In certain FSM states, data from the TDI port is captured. From the idle state, by driving the TMS signal in the appropriate way, the user chooses to enter either the select data or the select instruction state and then the data/instruction capture state. From that point on, the procedure is identical for both instruction and data capturing. The only difference is that the captured values are stored in different registers. In the data/instruction shift state, values from the TDI port are stored in a shift register as long as the FSM stays in that state. Afterwards, the FSM enters the update data/instruction state and then returns to the idle state, with the complete data/instruction value stored in the appropriate register. A block diagram of the complete JTAG FSM with all the state changes and the TMS signal values is shown in Fig. 1.

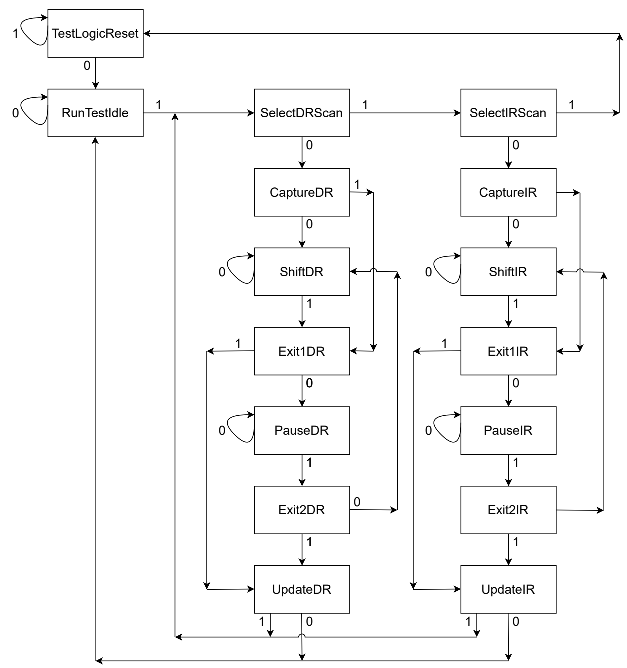


Fig. 1. A block diagram of the complete JTAG finite-state machine with all the state transitions and the TMS signal values.

B. AXI4 and TileLink Interfaces

Advanced eXtensible Interface 4 (AXI4) protocol is a parallel, synchronous, high-frequency multi-master and multi-slave communication interface. It is tailored mainly for the on-chip communication, which makes it suitable for the systems mentioned above. AXI4 interfaces consists of a vast number of different signals, with many of them optional, making it a versatile interface applicable to various different systems and applications. Even though it is described as multi-master and multi-slave interface, in every transaction only a single master and single slave communicate with each other. All the masters and the slaves are mutually connected through an interconnect bus. In the JTAG to memory-mapped bus master bridge module, this interface is used as a master interface and in most applications that include it, because of its sole purpose, the only active master is the module itself, whereas one or more slaves could exist. A simplified block diagram of an example of such a system is shown in Fig. 2. AXI4 interface protocol consists of 5 different channels: write address (AW), read address (AR), write data (W), read data (R) and response (B). Every one of those channels work on the handshaking principles and thus contain the pair of ready/valid signals. All the channels except the response channel have data signal among the various others signal whose function is to additionally describe and secure successful data transactions. Apart from the single read and single write data transactions, AXI4 interface supports the burst read and burst write data

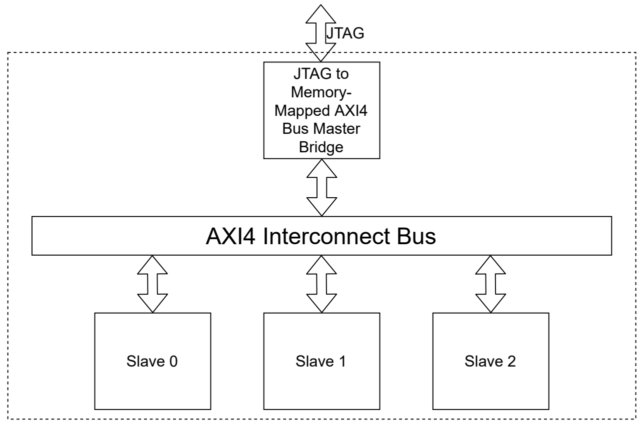


Fig. 2. A simplified block diagram of a system with a JTAG to memory-mapped bus master bridge and three slave peripherals attached to an AXI4 interconnect bus.

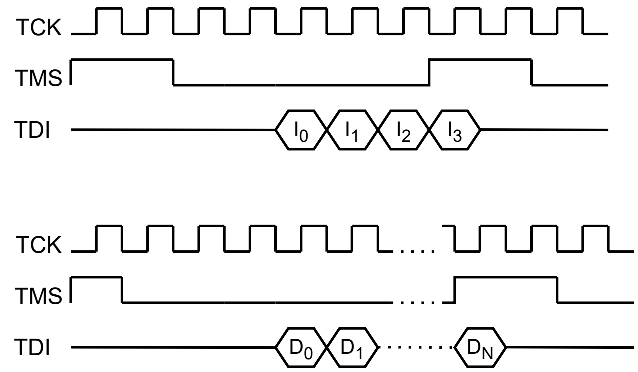


Fig. 3. JTAG input signal diagrams for correctly sending an instruction code (upper diagram) or a data value (bottom diagram) when the JTAG finite-state machine is in the idle state.

transactions.

TileLink is a parallel, synchronous, high-frequency multi-master and multi-slave communication protocol, in some extent similar to the AXI4 protocol. It is also designed mostly for the on-chip communication, with a special emphasis on the cache coherence transactions. The communication between a master and a slave (sometimes called a client and a manager in terms of this interface) is also performed based on the handshaking protocol. Overall, five communication channels can exist, with only two mandatory: channels A and D are mandatory, while channels B, C and E are optional. Each of the channels consists of several signals, including a data signal, a couple of ready/valid signals and some other signals used to describe and control transactions. The mandatory channel A flows from master interface to slave interface, carrying request messages sent to a particular address. Then, the slave responds to the master's request through the mandatory channel D. Other channels B, C and E are optional, and they are utilized for complete TileLink cached protocol, where channels B and C have similar functions as channels A and D respectively, whereas channel E is used as a final acknowledgment channel. In the module from the topic of this paper, however, only mandatory channels are used, and therefore, not much attention is provided to other three channels. TileLink protocol interface can be used as the master interface of the module instead of the AXI4 interface.

C. Defined Instructions

As it can be concluded from the previous paragraphs and sections, the JTAG to memory-mapped bus master bridge is a module that integrates two different interfaces with its dedicated controllers: JTAG and AXI4/TileLink. Those two controllers operate independently, they are even driven by different clock signals (TCK for JTAG controller and system clock signal for the AXI4/TileLink controller), but they are mutually synchronized through the internal signals. The JTAG controller works on the basis of the previously described JTAG FSM, while AXI4/TileLink has its own FSM. The whole module has the following, rather simple, signal flow. User calls an instruction by driving the serial JTAG input signals. The JTAG controller accepts the data, converts it to the parallel form and sends it to the AXI4/TileLink controller who, if it is recognized as a write instructions or a read instruction or any subvariant of them, initiates the transaction between the module and the appropriate slave. JTAG input signal values for correctly sending an instruction code or a data value when the JTAG FSM is in the idle state are shown on the timing diagrams in Fig. 3. In order to send the data value to the serial TDI input correctly, the least significant bit of the data should be sent first. It is strongly recommended that, prior to using the module, the user ensures that the JTAG FSM enters the reset state. It is achieved by driving the TMS signal with the active high value for five straight TCK cycles. By doing so, no matter what state was the JTAG FSM in, it will enter the desired reset state.

Total of four types of data transactions can be initiated by the AXI4/TileLink master interface on the interconnect bus: write, read, burst write and burst read. Also, nine instructions that the user can call through the JTAG interface are defined. The purpose of them is to enable the data transactions to be performed and to pass all the information needed for it. Each instruction can require either both instruction code and data itself to be provided, or just the instruction code. Both of them are captured in their dedicated JTAG FSM state. After the input serial data arrived entirely to the JTAG controller, it sends both the instruction code and the data to the AXI4/TileLink controller. A list of defined instructions, along with their codes and descriptions, is the following:

- 0x01 - Write instruction, initiates the AXI4/TL controller to begin writing the acquired data to the acquired address.
- 0x02 - Address acquire instruction, accepts the serial data as the address for the read/write instruction.
- 0x03 - Data acquire instruction, accepts the serial data as the data for the read/write instruction.
- 0x04 - Read instruction, initiates the AXI4/TL controller

to begin reading data from the acquired address.

- 0x08 - Number of burst transactions acquire instruction, accepts the serial data as the number of the read/write instructions during one burst transfer cycle.
- 0x09 - Burst write instruction, initiates the AXI4/TL controller to begin performing acquired number of the write transactions. Data is written to the consecutive addresses.
- 0x0A - Data index number acquire instruction, accepts the serial data as the index number of data to be acquired using the following instruction for the burst read/write transfer.
- 0x0B - Indexed data acquire instruction, accepts the serial data as the data at the acquired index number for the burst read/write transfer.
- 0x0C - Burst read instruction, initiates the AXI4/TL controller to begin performing acquired number of the read transactions. Data is read from the consecutive addresses.

Before the write instruction, both the address acquire and the data acquire instructions must be performed. Before the read instruction, the address acquire instruction must be performed. For the burst write instruction, the data for every single transaction must be acquired beforehand, as well as the total number of burst transactions for both the burst write and the burst read instructions. Two read/write/burst read/burst write instructions of the same type cannot appear sequentially one right after another, there must be at least one other instruction between the two. After performing the read or burst read instruction, read data appear on the serial output JTAG TDO data port, with the TDO driven signal having the active high value. All the instruction codes that are not mentioned in this paper can be assumed to be the no-operation (NOP) instructions.

III. A Design Generator and Its Implementation

Previously depicted JTAG to memory-mapped bus master bridge have been captured inside Chisel 3 hardware design generator. Both solely and in a combination with numerous slave peripheral modules, the generator has been thoroughly tested using standard Chisel verification and implementation paths for FPGA design flow. The design generator is made available [9] for public use as a free and open-source hardware library.

The generated module itself consists of two main subparts: JTAG controller and AXI4/TileLink controller, with their interfaces and internal communication signals. A block diagram of the JTAG to memory-mapped bus master bridge with all its interfaces and two main submodules is depicted in Fig. 4.

The JTAG controller is the submodule that communicates with the user. Its main purpose is to accept the serial data from the JTAG user interface, pack it in the appropriate format and send it to the AXI4/TileLink controller, as well as to accept the parallel data read from the slave peripherals and to put it on the JTAG serial data output. The JTAG controller is based on the previously depicted, standard JTAG FSM. Besides the common JTAG interface signals TCK, TMS, TDI and TDO, several more I/O signals exist. To begin with, TDO signal is divided into the one-bit-wide output signals: TDO data, which represents the serial output data, and TDO driven, which serves as the data valid signal. Those two signals are active at the falling edge of the TCK clock. There is also an asynchronous reset input signal which transits the JTAG FSM current state to the reset state.

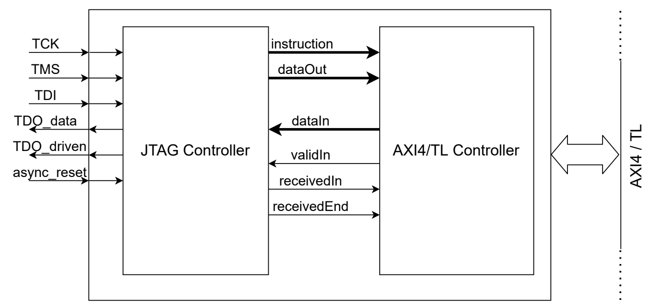


Fig. 4. A block diagram of the JTAG to memory-mapped bus master bridge with all its interfaces and the two main submodules.

Through the TDI JTAG serial input pin, the user can send either an instruction code or the data value itself. The JTAG can distinguish between those two thanks to the TMS control signals. Arrived data is translated from serial to parallel data format by using two shift registers, one for both instruction and data values. Data from the shift registers are stored into the two data registers, one at a time, when the JTAG FSM enters the appropriate data/instruction capture states. Data from those two registers are sent separately to the AXI4/TileLink controller.

A. The AXI4 and TileLink Controllers

The AXI4 and TileLink controllers are similar to one another, with the obvious difference in the master interface signals. They accept the instruction and data values from the JTAG controller, recognize the instruction code and take the action correspondingly. If the instruction code suits the either read or write instruction code, the controller initiates the communication with the appropriate slave peripheral through the AXI4/TileLink interconnect bus. The instruction code and data values are stored into the two separate registers. The value from the instruction register is constantly checked and compared to the instruction code of each of the four data transfer instructions (write, read, burst write, burst read). When those two values match, the appropriate flag value is set to the active high (a flag exist for every one of those four instructions) and that signifies that the appropriate instruction is set to be executed. Apart from signalizing that the instruction should be performed, the flag signals are used to prevent other instructions to be executed until the end of the current instruction. For the purpose of the burst transfers, a counter that counts the number of performed transfers is implemented inside the controller. Both controllers rely on their own FSMs which secures the correctness of the communication with the peripherals.

The AXI4 and TileLink controllers also send the data read from the peripherals to the JTAG controller. For that purpose, several more internal signals exist. Apart from the one that carries data values to the JTAG controller, there are signal that marks the validity of the arrived data and two signals that mark that the JTAG controller has received the data and that the all bits of the data were sent to the output TDO pin.

The AXI4 controller FSM has the task to control the communication with the slave peripherals. State transitions are realized thanks to either flag values mentioned above, or the AXI4 signal values from the slave, such as ready signal for the handshaking protocol. Following states exist:

- `sIdle` - The idle state, the FSM stays in this state until the write instruction or the read instruction flag is set.
- `sSetDataAndAddress` - The state in which address is set on the AW channel, data is set on the W channel and valid signals are set on both the AW and W channels. The FSM stays in this state until the ready signals are not set on both the W and AW channels or until a counter which ensures that the FSM isn't stuck in this state counts out.
- `sResetCounterW` - The state in which the mentioned counter is reset. Stays in this state for exactly one clock cycle.
- `sSetReadyB` - The state in which the ready signal is set on the acknowledgement B channel. The FSM stays in this state until the B channel valid signal is not set or until the counter which ensures that the FSM is not stuck in this state counts out. The write instruction flag is reset in this state.
- `sSetReadAddress` - The state in which the address and the valid signals are set on the AR channel. The FSM stays in this state until the AR channel ready signal is not set or until the counter which ensures that the FSM is not stuck in this state counts out.
- `sResetCounterR` - The state in which the mentioned counter is reset. The FSM stays in this state for exactly one clock cycle.
- `sSetReadyR` - The state in which the ready signal is set on the R channel and data is read from the same channel. The FSM stays in this state until the R channel valid signal is not set or until the counter which ensures that the FSM is not stuck in this state counts out.
- `sDataForward` - The state in which the read data is forwarded to the JTAG controller, along with the active valid signal. The FSM stays in this state until the JTAG controller does not confirm that the data is received. The read instruction flag is reset in this state.

A state transition diagram for the AXI4 FSM is depicted in Fig. 5. Note that the states for the write and the read instructions only are shown. The reason for the deficiency of the other two is simply the similarity to the depicted ones. The only novelty for the burst transfers is the fact that after the completed single data transfer, FSM enters `sIdle` state only if the burst transfers counter has counted out. Otherwise, the AXI4 FSM enters `sSetDataAndAddress/sSetReadAddress` state to perform another transfer.

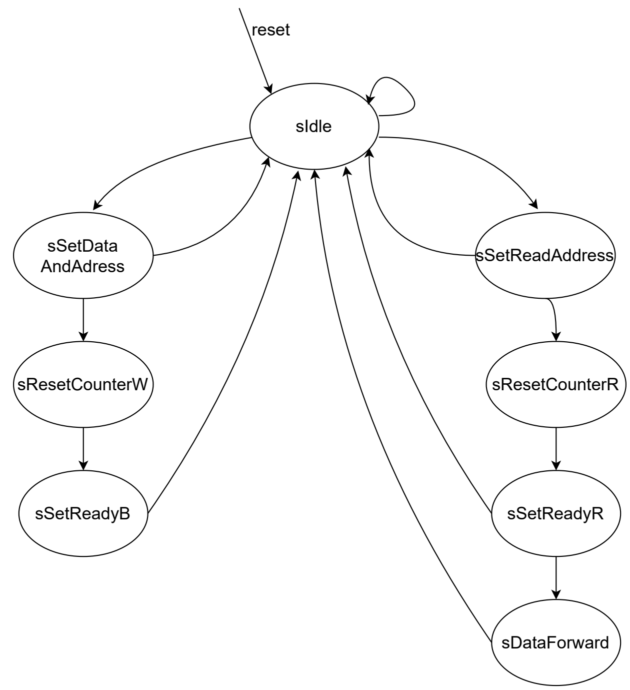


Fig. 5. A state transition diagram for the read and the write instructions of the AXI4 FSM.

The TileLink controller FSM has the same task as the AXI4 FSM. Its state transitions are also realized thanks to either the flag values or the signal values received from the slave peripheral. Although the states themselves are not identical to the ones from the AXI4 FSM, mostly because of the differences between the interfaces, the overall principles are the same. Therefore, they will not be elaborated in this paper.

B. The Chisel Generator

The Chisel generator of the JTAG to memory-mapped bus master bridge has few parameters that can impact the characteristics of the generated instances. Data and address buses widths for all the AXI4/TileLink channels can differ between 32 and 64. The instruction code width is also changeable. As the current number of instructions is nine, the width of four bits is sufficient for all the instruction codes. Another parameter represents the code for the initial instruction. It is strongly recommended that any code of the NOP instruction is provided as this parameter. Maximum number of transfers in a burst cycle can also take different values, as well as the set of the addresses that the module's master interface can access. Even though the Chisel language is extremely suitable for parameterization of the modules, this capability is not exploited a lot in this case due to the nature of the proposed module itself.

For the implementation of the generator, several exploited open-source Chisel libraries worth mentioning exist. Chipsalliance's Rocketchip library [10] is extensively used. It provided the extremely valuable classes for the implementation of both AXI4 and TileLink master interfaces, as well as of the interconnect bus and memory-mapped address space. Also, the Ucb-art's Chisel-JTAG library [11] was beneficial for the realization of the module. Its JTAG FSM design with some other modules, such as shift registers and I/O bundles, were utilized. The generator itself was integrated into the Ucb-bar's Dsptools library [12].

IV. Implementation and Verification Results

There are several stages of the JTAG to memory-mapped bus master bridge testing. The first one represents the usage of the software simulations. For the performance of these tests, Chisel testers are utilized to drive the JTAG input signals. Apart from testing the module solely, it was also verified experimentally using various other modules with memory-mapped control and status registers, from the simple ones, such as a streaming multiplexer, to more complicated ones, such as a parameterizable numerically-controlled oscillator or run-time configurable fast Fourier transformation module. The tested module was also verified within the simulation environments with multiple slave modules.

Another stage of the proposed generator's verification is implementing and testing the generated instances on an FPGA-based development board. A Digilent's Arty A7 board with Xilinx Artix-7 FPGA family is used for it. All the generator's instances are synthesized for 100 MHz system clock frequency. The JTAG input signals were driven from the PC using the FTDI's C232HM-DDHSL-0 cable [13]. That cable contains the FT232H integrated circuit [14] and represents the USB 2.0 hi-speed to multi-protocol synchronous serial engine (MPSSE) cable. For the utilization of the FT232H chip and the cable itself, Pyftdi open-source library [15] is used. The JTAG input signals are generated as the general purpose input/output (GPIO) signals. JTAG clock frequency was set to 15 MHz (GPIO pins can work with the frequency up to 30 MHz) which is the convenient speed having in mind that the JTAG clock frequency is obligated to be lower than the system clock frequency in order for the module to work properly. Similar to the verification using software simulations, several additional modules with the memory-mapped registers were used to validate the functional correctness of the proposed generator's module.

The FPGA resource utilization for the JTAG to memory-mapped bus master bridge is not significant due to the lack of the complex arithmetic or logic operations and few used registers. Moreover, it is expected for the peripheral's utilized resources to be drastically more numerous.

V. Conclusion

In this paper, a generator of the JTAG to memory-mapped bus master bridge implemented using Chisel hardware design language is proposed. Mainly, it is used for configuring, testing and debugging the peripheral modules with memory-mapped input/output interface in the systems without the processing core or where the processing core is set to remain inactive in terms of communication with the slave peripherals through the interconnection bus. Even tough it seems to be a complete product right now, theoretically speaking a lot of space was saved for the further upgrade, primarily to the Chisel's parameterizable characteristics.

The generated instances of the JTAG to memory-mapped bus master bridge were tested and verified by both using software simulations and mapping onto a commercial FPGA development board. Numerous additional modules with the memory-mapped slave interface are utilized in the testing process. The module from the topic of this paper proved to be trustworthy regarding its functionality and performance.

Acknowledgements

The authors would like to thank NOVELIC d.o.o. for financially and logistically supporting the work on this project.

References

[1] N. Y. Song, Y. J. Yu, W. Shin, H. Eom, and H. Y. Yeom, "Low-latency memory-mapped I/O for data-intensive applications on fast storage devices," in *2012 SC Companion: High Performance Computing, Networking Storage and Analysis*, 2012, pp. 766–770.

[2] *Intel 64 and IA-32 Architectures Software Developer's Manual*, Order number: 325383-060us ed., Intel, September 2016.

[3] *AMBA AXI and ACE Protocol Specification*, Arm ihi 0022e (id033013) ed., ARM, February 2013.

[4] *SiFive TileLink Specification*, Tilelink specification, version 1.8.0 ed., SiFive, Inc, August 2019.

[5] B. V. Ngo, P. Law, and A. Sparks, "Use of JTAG boundary-scan for testing electronic circuit boards and systems," in *2008 IEEE AUTOTEST-CON*, 2008, pp. 17–22.

[6] L. Ungar, H. Bleeker, J. McDermid, and H. Hulvershorn, "IEEE-1149.x standards: achievements vs. expectations," in *2001 IEEE Autotestcon Proceedings. IEEE Systems Readiness Technology Conference. (Cat. No.01CH37237)*, 2001, pp. 188–205.

[7] *JTAG to AXI Master v1.2*, Pg174 ed., Xilinx, February 2021.

[8] *Embedded Peripherals IP User Guide*, UG-01085 ed., Intel, March 2021, pages 56-60.

[9] V. D. Damnjanović and V. M. Milovanović, "JTAG to memory master (JTAG2MM) chisel generator," www.github.com/milovanovic/jtag2mm, accessed: 2021/06/21.

[10] K. Asanović, R. Avizienis, J. Bachrach, S. Beamer, D. Biancolin, C. Celio, H. Cook, D. Dabbelt, J. Hauser, A. Izraelevitz, S. Karandikar, B. Keller, D. Kim, J. Koenig, Y. Lee, E. Love, M. Maas, A. Magyar, H. Mao, M. Moreto, A. Ou, D. A. Patterson, B. Richards, C. Schmidt, S. Twigg, H. Vo, and A. Waterman, "The rocket chip generator," EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2016-17, Apr 2016. [Online]. Available: http://www2.eecs.berkeley.edu/Pubs/TechRpts/2016/EECS-2016-17.html

[11] C. M. Richard Lin, Paul Rigge, "chisel-jtag," www.github.com/ucb-art/chisel-jtag, accessed: 2021/06/21.

[12] P. R. Chick Markley, Angie Wang, "dsptools," www.github.com/ucb-bar/dsptools, accessed: 2021/06/21.

[13] *C232HM*, version 1.3 ed., FTDI, document No.: FT000401 Clearance No.: FTDI 214.

[14] *FT232H*, version 2.0 ed., FTDI, document No.: FT000288 Clearance No.: FTDI 199.

[15] E. Blot, "PyFTDI," www.eblot.github.io/pyftdi/, accessed: 2021/06/21.

Allpass Based Double Notch IIR Filters with Constant Phase

Goran Stančić, Ivana Kostić and Stevica Cvetković

***Abstract*—Narrow stopband filters with two notch frequencies and piecewise constant phase are investigated in this paper. The notch filters are determined by allpass subfilter phase approximation. Obtained filters with simple and double poles are compared in conditions when fractional part of the coefficients is represented with limited number of bits.**

***Index Terms*—Notch IIR filters, allpass filters, phase approximation, constant phase, quantization.**

I. Introduction

THE ideal notch filters have exactly zero magnitude at frequency which need to be removed from input signal spectrum and unity gain otherwise. In many devices in practice the power line frequency signal and corresponding harmonics are often treated as noise [1]. Over time, digital electronic components become faster and cheaper allowing designers to oversample input signal. In that case the neighboring harmonics start to go closer to one another at the frequency axis. Notch filters are part of radar systems, control and instrumentation systems, medical applications and communications systems. In order to keep the distortion of desired signal as low as possible, the stopband of the notch filter should be as narrow as possible. In this paper problems associated with close notch frequencies will be observed. All of the presented results are given for double notch filters but it is easy to modify proposed method for arbitrary number of notch frequencies.

II. Realization structure

In addition to standard realization structures filters can be obtained by parallel connection of two allpass filters [2]. The magnitude of resulting filter depends on phases of applied allpass filters. That is a reason why design of the linear/constant phase IIR filters comes down to the allpass phase approximation problem.

In practice linear and constant phase are ultimate goals to avoid phase distortions [3]. To obtain the notch filter with approximately constant phase one allpass filter becomes direct path as shown in Fig. 1 [4]. The notch filter with linear phase will be achieved if one allpass filter is pure delay [5].

Goran Stančić, Ivana Kostić and Stevica Cvetković are with University of Niš, Faculty of Electric Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: ivana.kostic@elfak.ni.ac.rs

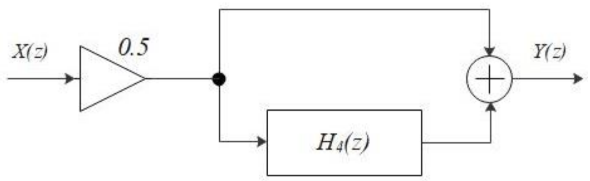


Fig. 1. Double notch filter realised as parallel connection of direct path and fourth order allpass filter.

The transfer function of constant phase notch filter is

$$H(z) = 0.5(1 + H_4(z)) \tag{1}$$

where $H_4(z)$ represents transfer function of allpass filter of the form

$$H_4(z) = \prod_{i=1}^{2} \frac{(\rho_i - e^{-j\theta} z^{-1})(\rho_i - e^{j\theta} z^{-1})}{(1 - \rho_i e^{j\theta} z^{-1})(1 - \rho_i e^{-j\theta} z^{-1})} \tag{2}$$

taking into account the fact that allpass filters have conjugate-reciprocal pole-zero pairs. Magnitude of the notch filter directly depends on the allpass filters phase φ

$$\left|H(e^{j\omega})\right| = \left|cos\frac{\varphi(\boldsymbol{\rho}, \boldsymbol{\theta}, \omega)}{2}\right| \tag{3}$$

where $\boldsymbol{\rho}$ and $\boldsymbol{\theta}$ represent moduli and phase angles of the allpass filters poles, respectively. Every pole and zero contribute to the phase with π/2 radians making fourth order filter to reach -4π radians phase at Nyquist frequency as shown in Fig. 2.

The closer a pole is to the unit circle the higher negative slope is at frequencies in vicinity of pole position. The phase is monotonically decreasing function of frequency with emphasized jump around pole position. Fourth order transfer function also could be obtained with two simple poles. This case is marked with d) in Fig. 2. Now poles are not at the same frequency and two separate phase jumps of approximately -2π radians could be observed.

According to (3), filter realized with described parallel structure possess passbands at frequencies ω where $\varphi(\boldsymbol{\rho}, \boldsymbol{\theta}, \omega)$ approximates $2k\pi$ with allowed tolerance ε, for $k \in \mathbb{Z}$. Stopbands would be obtained at frequencies where phase value is approximately $2(k+1)\pi$.

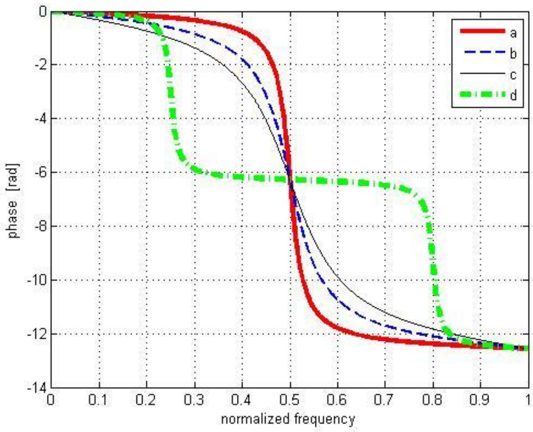


Fig. 2. Phase of allpass filter of fourth order for double pole moduli a) ρ=0.94, b) ρ=0.85, c) ρ=0.77 for θ=0.5π and d) simple poles $\rho_1 = 0.968, \rho_2 = 0.967$ for $\theta_1 = 0.25\pi$ and $\theta_2 = 0.8\pi$.

For predefined attenuation a in decibels, at stopbands or passbands boundary frequencies, allowed phase approximation error has value

$$\varepsilon = 2\mathrm{arccos}\,(10^{-a/20}) \tag{4}$$

III. DESIGN PROCEDURE

If filter with two notch frequencies has double pole, to determine the transfer function only two unknown values need to be calculated- modulus and angle of the pole. The positions of notch frequencies are of highest importance, so one need to solve next system of equations

$$\begin{gathered}\varphi(\rho,\theta,\omega_{n1}) = -\pi \\ \varphi(\rho,\theta,\omega_{n2}) = -3\pi.\end{gathered} \tag{5}$$

Notch filter with two simple poles has four unknown parameters. It allows two boundary frequencies to be controled. System of equations that provide notch filter transfer function has now form

$$\begin{gathered}\varphi(\boldsymbol{\rho},\boldsymbol{\theta},\omega_{n1}) = -\pi \\ \varphi(\boldsymbol{\rho},\boldsymbol{\theta},\omega_{n2}) = -3\pi \\ \varphi(\boldsymbol{\rho},\boldsymbol{\theta},\omega_{l1}) = -\varepsilon \\ \varphi(\boldsymbol{\rho},\boldsymbol{\theta},\omega_{r3}) = -4\pi+\varepsilon\end{gathered} \tag{6}$$

where ω_{l1} and ω_{r3} represents boundary frequencies of the first and the third passband, respectively.

Instead of passband edges, it is possible to use stopband edges in system (6) but in (4) minimal attenuation in stopband need to be applied, with minimal modifications of last two equations in (6) ($-\varepsilon$ will be changed with $-\pi \pm \varepsilon$ and $-4\pi + \varepsilon$ with $-3\pi \pm \varepsilon$).

Systems of equations (5) and (6) could be solved applying some iterative procedure which demands the initial solution. System (6) is given in alternative form

$$\boldsymbol{A\Delta} = \boldsymbol{B} \tag{7}$$

after approximating phase $\varphi(\boldsymbol{\rho},\boldsymbol{\theta},\omega)$ by truncated Taylor series, where elements of matrix $\boldsymbol{A}$ are

$$a_{ij} = \begin{cases}\dfrac{d\varphi(\omega_i)}{d\rho_j}, & j = 1,2 \quad i = 1,..,4 \\ \dfrac{d\varphi(\omega_i)}{d\theta_{j-2}}, & j = 3,4 \quad i = 1,..,4\end{cases} \tag{8}$$

elements of column vector $\boldsymbol{B}$ are

$$\boldsymbol{B} = \begin{bmatrix} -\pi - \varphi(\boldsymbol{\rho}^*,\boldsymbol{\theta}^*,\omega_{n1}) \\ -3\pi - \varphi(\boldsymbol{\rho}^*,\boldsymbol{\theta}^*,\omega_{n2}) \\ -\varepsilon - \varphi(\boldsymbol{\rho}^*,\boldsymbol{\theta}^*,\omega_{l1}) \\ -4\pi + \varepsilon - \varphi(\boldsymbol{\rho}^*,\boldsymbol{\theta}^*,\omega_{r3}) \end{bmatrix} \tag{9}$$

and vector of increments to be found is given with

$$\boldsymbol{\Delta} = [\Delta\rho_1;\ \Delta\rho_2;\ \Delta\theta_1;\ \Delta\theta_2] \tag{10}$$

In every iterative step system (7) is solved, the modulus and phase angle of poles are corrected until maximal absolute value of elements of column vector $\boldsymbol{\Delta}$ becomes less than predefined small value (in all given examples 10^{-10} is applied). As a good initial solution one could choose values

$$\rho^* = 0.9 \quad \text{and} \quad \theta^* = \frac{\omega_{n1}+\omega_{n2}}{2} \tag{11}$$

for double pole and

$$\boldsymbol{\rho}^* = [0.9; 0.9] \quad \text{and} \quad \boldsymbol{\theta}^* = [\omega_{n1};\ \omega_{n2}] \tag{12}$$

for simple poles. Extensive experiments shown that the final solution would be reached in less than ten iterations for arbitrary feasible input parameters.

IV. RESULTS

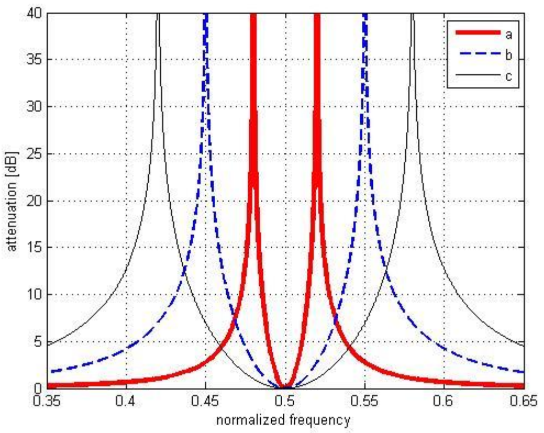


Fig. 3. The filters with two notch frequencies realized with double pole for a) $\omega_{n1} = 0.48\pi, \omega_{n2} = 0.52\pi$, b) $\omega_{n1} = 0.45\pi, \omega_{n2} = 0.55\pi$ and c) $\omega_{n1} = 0.42\pi, \omega_{n2} = 0.58\pi$.

In Fig. 3 are displayed characteristics of double notch filter attenuation for different zero magnitude frequencies. All filters have a double pole. Taking into account the phase characteristics shown in Fig. 2, to achieve more distance between the notches it is inevitable to move the pole closer to the origin. That is good for the stability because the pole moves further from the unit circle. Lower pole modulus values provoke lower phase slope, so the transition zones and stopbands become wider at the expense of passbands. This feature points to fact that double pole notch filter has restricted application. It is not possible to choose higher order allpass filter in attempt to improve notch filters characteristics, as in case of linear phase filters.

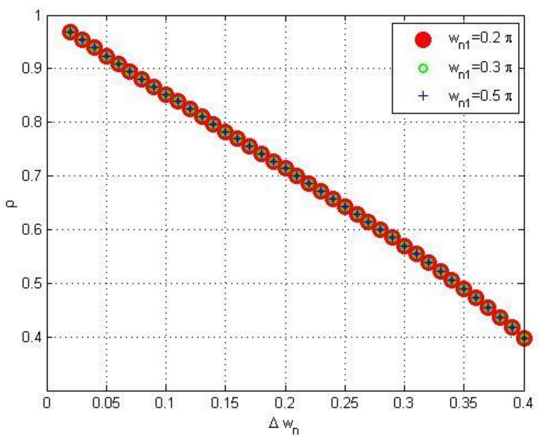


Fig. 4. The dependance of double pole modulus on notch frequencies gap for different ω_{n1} locations ($\Delta\omega_n = \omega_{n2} - \omega_{n1}$).

From Fig. 4 could be observed that most significant influence on double pole modulus has the notch frequencies gap. The very value of notch frequencies location have no visible impact. In practical realization of digital filter, the number of bits for filters coefficients representation need to be defined.

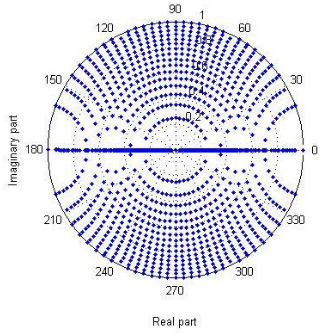


Fig. 5. Possible positions of filters poles for fixed point arithmetics when 4 bits are reserved for fractional part of transfer function coefficients.

Finite number of bits leads to rounded values of coefficients so realized filter characteristics just approximate derived ones. Possible positions of poles of the second order transfer function are displayed in Fig. 5 in case four bits are dedicated to fractional parts. In other words, calculated transfer function will be replaced with approximated one and obtained poles have to move from obtained positions to available locations like in Fig. 5.

The Fig. 5 indicates the fact that one can expect bigger error as consequence of quantization if notches are positioned at low and high frequencies.

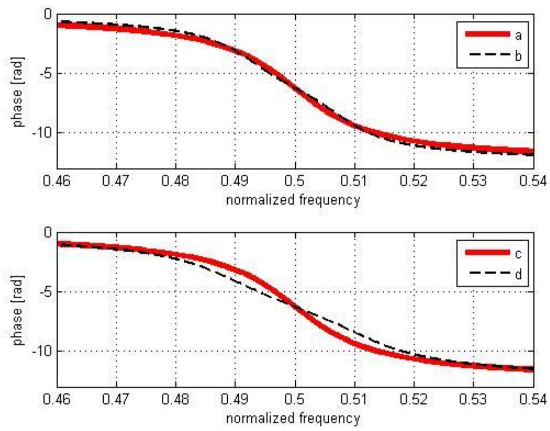


Fig. 6. Phase of notch filters ($\omega_{n1} = 0.49\pi$, $\omega_{n2} = 0.51\pi$) with a),c) double and b), d) simple poles before and after quantization fractional parts with 4 bits, respectively.

As first example filters with $\omega_{n1} = 0.49\pi$ and $\omega_{n2} = 0.51\pi$ are designed with desribed procedure. Attenuation of 1 dB is chossen in passbands. Boundary frequencies are ω_{l1}=0.47 π and ω_{r3}=0.53 π. Corresponding phase, attenuation and poles location are presented in Fig. 6, Fig. 7 and Fig. 8, respectively.

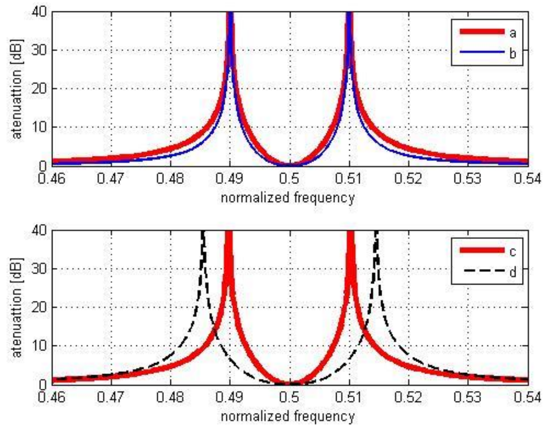


Fig. 7. Attenuation of notch filters ($\omega_{n1} = 0.49\pi$, $\omega_{n2} = 0.51\pi$) with a), c) double and b), d) simple poles before and after quantization fractional parts with 4 bits, respectively.

As it was expected, the phase undergo changes as repercussion of quantization, causing notch frequencies to displace. The notches are misplaced for 0.0142 for simple poles and $9 \cdot 10^{-4}$ for double pole case. Symmetry helps double pole filter less to degrade. The reason can be found in Fig. 5, where one can observe that pole with θ=0.5π will change only pole modulus as given in Fig. 8. Simple poles changed both moduli and phase angles causing significant mismatch between desired and obtained notches.

All obtained filters are realized as serially-cascaded second-order sections. Denominator coefficients of second order sections of the allpass filter with simple poles are given in Table I. All presented results are obtained in Matlab. The coefficients of allpass filter with double pole are presented in Table II.

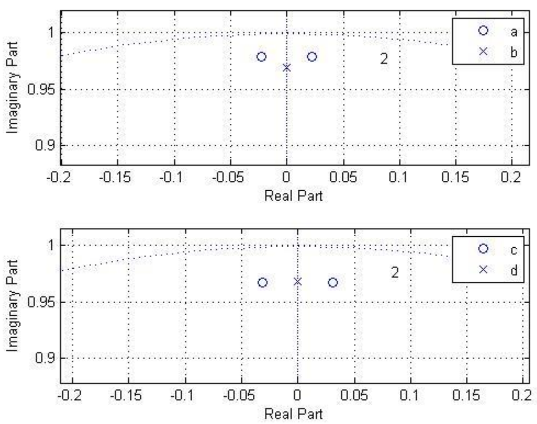


Fig. 8. Location of notch filters poles ($\omega_{n1} = 0.49\pi$, $\omega_{n2} = 0.51\pi$) with a), c) simple and b), d) double poles before and after quantization fractional parts with 4 bits, respectively.

TABLE I
COEFFICIENTS OF THE SECOND ORDER SECTIONS (SIMPLE POLES)

1.	0.0451	0.9581
1.	-0.0451	0.9581

TABLE II
COEFFICIENTS OF THE SECOND ORDER SECTIONS (DOUBLE POLE)

1.	0.	0.9391
1.	0.	0.9391

After quantization, with four bits dedicated to the fractional part, new values for second order sections coefficients are obtained as given in Table III. Table IV contains coefficients of allpass filter with a double pole. Because of existing symmetry second order sections have one coefficient equal to zero demanding less multipliers and adders in hardware realization.

TABLE III
COEFFICIENTS OF THE SECOND ORDER SECTIONS AFTER QUANTIZATION (SIMPLE POLES)

1.	0.0625	0.9375
1.	-0.0625	0.9375

TABLE IV
COEFFICIENTS OF THE SECOND ORDER SECTIONS AFTER QUANTIZATION (DOUBLE POLE)

1.	0.	0.9375
1.	0.	0.9375

For second example filters with $\omega_{n1} = 0.10\pi$ and $\omega_{n2} = 0.11\pi$ are chosen. For design procedure values ω_{l1}=0.09 π, ω_{r3}=0.12 π and α =3 dB are adopted. Obtained phase, attenuation and poles location are presented in Fig. 9, Fig. 10 and Fig. 11, respectively. These filters have poles in area where possible pole locations are scattered. As consequence, quantization of filter coefficients will seriously degrade characteristics. Close notches demand poles to be near the unit circle to provide enough steep slope.

Even 5 bits dedicated to the fractional part was not enough to stop a pole at the end of the unit circle. From (2) it is obvious that in such a case influence of the allpass zero and pole is identical, forcing the transfer function to degrade to second order. As a result, filter possess only one notch, as given in Fig. 10 d). The double pole filter has two notches after quantization but rear possible pole positions considerably influence the notches to displace.

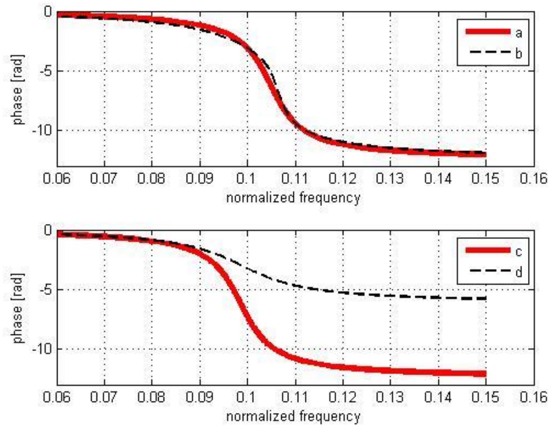


Fig. 9. Phase of notch filters ($\omega_{n1} = 0.10\pi$, $\omega_{n2} = 0.11\pi$) with a),c) double and b), d) simple poles before and after quantization fractional parts with 5 bits,respectively.

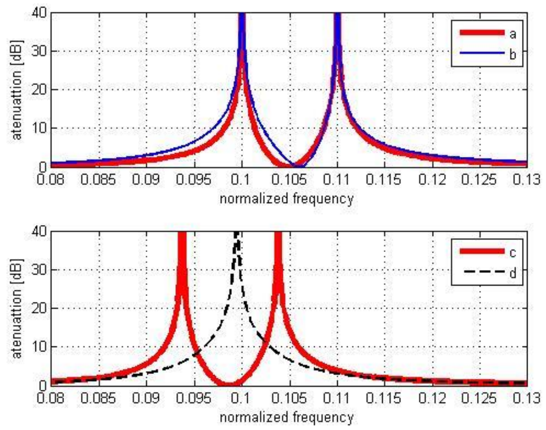


Fig. 10. Attenuation of notch filters ($\omega_{n1} = 0.10\pi$, $\omega_{n2} = 0.11\pi$) with a), c) double and b), d) simple poles before and after quantization of fractional parts with 5 bits, respectively.

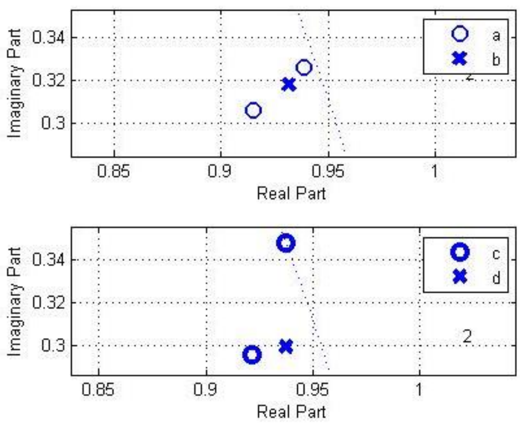


Fig. 11. Location of notch filters poles ($\omega_{n1} = 0.10\pi$, $\omega_{n2} = 0.11\pi$) with a), c) simple and b), d) double poles before and after quantization fractional parts with 5 bits, respectively.

V. Conclusion

Double notch filters with constant phase in passbands are investigated in this paper. Two similar solutions are compared and impact of quantization is analyzed. Parallel allpass structure guarantee low passband sensitivity. Quantization effects primarily affect notch filters stopband, moving away locations of notches from desired positions. The double pole filters are not good choice in case when gap between notches is wider than 0.2π because low phase slope causes transition zones to spread, degrading selectivity. On the other hand, the pole has lower modulus if filter possess double pole, what is guarantee to remain stable after quantization and still to have both notches. The distance of simple pole from the unit circle is always smaller compared to double pole. As consequence, after quantization if notches are close to each other it may occur one or both simple poles to finish at the unit circle and quantized version of filter lose selectivity. Design of numerous notch filters with different notches location have been shown that double pole solution is better option for close notches and small number of bits dedicated to the coefficients fractional part.

Acknowledgment

This work has been supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

References

[1] P. Zahradnik and M. Vlček, “Notch filtering suitable for real time removal of power line interference”, *Radioengineering*, vol. *22*, no. *1*, pp. *186-193*, 2013.

[2] P. P. Vaiduanathan P. A. Regalia and S. K. Mitra, “A new approach to the realization of low sensitivity IIR digital filters”, *IEEE Trans. on Acoust., Speech and Sig. Proc.*, ASSP-34, pp. *350-361*, 1986.

[3] S. C. Dutta Roy, B. Kumar and S. B. Jain, “FIR notch filter design – A review”, *Facta Universitatis ser. Electronic and Energetics*, vol. *14*, no. *3*, pp. *295-327*, 2001.

[4] G. Stančić and S. Nikolić, “Design of narrow stopband recursive digital filters”, *Facta Universitatis ser. Electronic and Energetics*, vol. *24*, no. *1*, pp. *119-130*, 2011.

[5] S. Nikolić , G. Stančić and S. Cvetković, “Design of nearly linear phase double notch digital filters with close notch frequencies”, *IET Signal Processing*, vol. *12*, no. *9*, pp. *1107-1114*, 2018.

Free/Open Source EDA Tools Application in Digital IC Design Curricula

Aleksandar Pajkanovic, *Member*, *IEEE*

***Abstract*—This paper represents a report on how free and open source software EDA tools may be used to organize a digital integrated circuits design course at the university level, without any financial investments in licenses. The course is built around several publicly available processor cores. These are of different complexity so first intrinsic properties are investigated. Then, using more complicated designs we examine how to increase performance through pipeline and cache associativity configurations. In this way we introduce RISC-V ISA and Chisel into the curricula. Finally, we provide a short overview of tools for automated design, from RTL all the way to silicon.**

***Index Terms*—open source, digital design, RISC-V, Chisel**

I. Introduction

Since the first implementation of a transistor was demonstrated at the Bell Labs back in mid-20th century [1], the industry has shifted from traditional, as established by the Industrial Revolution, to an economy based on the information technology (IT). That event, we understand today, represents the onset of the Information Age–the age characterized by rapid growth and development in all areas of life, driven by the semiconductor industry. Its workhorse, the CMOS technology process, is characterized by low power consumption, extreme scalability and ease of mass production. The ability to implement an idea, a solution using this technology, i.e. the ability to design integrated circuits (IC), or chips, has been of greatest importance for decades and it will be even more so in decades to come; namely, even though the Moore's Law [2] has ended–i.e. we do not advance our ICs by scaling anymore, but rather using advanced techniques [3] in design and verification phases, in order to achieve better performance–we still fabricate our solutions in silicon CMOS technology process.

Over the last five years, we have been establishing IC courses at the Faculty of Electrical Engineering in Banja Luka. This effort has been reported in [4], [5] and has recently culminated in fabrication ready circuits [6]. All this has been achieved using exclusively free and open source EDA software tools, with the help of many contributors - both students and tool developers. In those developments, however, main effort was in the analog domain, whereas in this particular paper we focus on digital IC design course, the examples it's built upon and tools used.

Since materials such as combinatorial and sequential circuits are covered in other courses, for this particular course we've decided to learn about more advanced concepts using the free intellectual property (IP) available in the community. Thus we study three processor designs, we learn a new approach to digital design (hardware construction, instead of description) and we drive the RTL code all the way to GDS, i.e. silicon ready file, using the free and open source toolchain.

In the next section we present motivation for writing up this paper, then we provide brief overviews of the RISC-V Instruction Set Architecture (ISA), *Chisel* - the hardware construction language (HCL) and processor cores we learn about and use to demonstrate theoretical concepts. Finally, in the sixth section, we present the free/open source digital IC design toolchain.

II. Motivation

There is no point in living during the Information Age, unless we are going to use its benefits. While boundaries are important parts of our lives and some should never be crossed, there are those boundaries that are not actually natural - these simply existed due to the fact that we knew not how to overcome them. This is not the matter of destroying those boundaries, but rather outgrowing them to improve the world and general quality of life. Such borders are those related to knowledge. With the Internet and its omnipresence - knowledge is omnipresent as well, for those who seek it. Our idea is to build the IC curricula by standing on the shoulders of the giants–of those who have had the chance to grow and develop for decades in this domain, thanks to another kind of boundaries. Therefore, we bring what's best on this planet right in our own court and thus enable our own students right here in Banja Luka to gain world-class expertise in the semiconductor industry, which, after all, is the most sophisticated commercially available technology process. And we do that without any financial investments - but rather simply: by reaching out.

Main motivation behind this paper is to contribute to the open source hardware community by sharing collected experiences and provide feedback on a subset of freely available tools and IPs, for all those who find themselves struggling to get started, at no cost, in this extremely interesting and exciting engineering and science area.

III. RISC-V

Computer architecture is, as most engineering areas nowadays are, an incredibly vast discipline. However, for the purposes of this short article, we will overly simplify and point out that it can be divided in two subdiciplines: software and

Aleksandar Pajkanovic is with the Faculty of Electrical Engineering, University of Banja Luka, Patre 5, 78 000 Banja Luka, The Serbian Republic, Bosnia and Herzegovina, e-mail: aleksandar.pajkanovic@etf.unibl.org

hardware. ISA is probably the most important interface in the universe as it serves to connect these two worlds. This has become clear with the IBM 360 appearance on the market, when the concept of ISA was first introduced. All ISAs with significant usage are proprietary, which does make sense when taking into account that the creators protect their intelectual property. However, *open* standards (the correct term would be open-source - to maintain analogy with software, where it all started, but there's no source code nor source files in this context, so we just call them open standards) like ethernet have proven successful. Successful, meaning that we have seen free-market competition through technical improvements, whence the end-users benefited the most [7]. Therefore, it is a crucial question to raise: why not create an open standard for the most important abstraction layer, the ISA?

In an answer, RISC-V (where V is a roman number five, thus pronounced) ISA has been designed - a completely free and open ISA, built and improved on the original RISC architectures. While it was designed originally for research and teaching, it is on its road to become a standard for industry implementations, as well. The final result is a simple and modular ISA, well suited to both high-performance systems and low-power embedded controllers. This is enabled by dividing the instructions into the obligatory base subset of ISA, present in any implementation, plus the optional extensions (subsets). The base is restricted to contain a minimum number of instructions sufficient for compilers, assemblers and linkers. If an operating system is to be used, an additional subset is required. Therefore, it is possible to look at the RISC-V as if it is actually a family of related ISAs. There are two base integer variants, RV32I and RV64I, each providing 32-bit or 64-bit width, respectively. The subset marked with E is designed for small microcontrollers, whereas the subset M enables multiplication and division, F is single-precision floating-pint, D for double-precision, etc. The most available RISC-V processors implement the RV64IMAFD flavor, which, due to its popularity, is marked with G (for general) [8]. Such is the rocket core, discussed in Section V of this paper.

RISC-V represents a perfect combination of simplicity and industrial application, and is, therefore, selected as the ISA to be studied in VLSI courses.

IV. CHISEL

Chisel is a hardware construction language, first introduced in [9], built as a domain specific language (DSL) embedded in Scala, with the idea to support advanced hardware design by providing important concepts established in software engineering. These include object orientation, functional programming, parameterized types and type inference. Such powerful abstraction features enable high level of code reuse, thus improving efficiency in constructing new hardware systems. In this way, while not claiming that Chisel is *better* in general, we do point out that Chisel does provide new paradigms in hardware design, that can increase productivity.

At first glance, looking at simple combinatorial module such as a multiplexer shown in Listing 1, there are no conceptual differences to standard Verilog HDL approach, aside from the syntax [10]:

```
class Mux2 extends Module {
  val io = IO(new Bundle {
    val sel = Input(UInt(1.W))
    val in0 = Input(UInt(1.W))
    val in1 = Input(UInt(1.W))
    val out = Input(UInt(1.W))
  })
  io.out := (io.sel & io.in1) |
            (~io.sel & io.in0)
}
```

Listing 1. Multiplexer 2/1 in Chisel

While some strangeness is present, due to inheritance and := operator, it is quite obvious we have a group of input/output ports packed in a `struct`-like piece of code (a *Bundle*) and some wiring in the last line. Looking at state elements, Chisel is still quite comparable to Verilog, as demonstrated by the 4-bit shift register in Listing 2 [10]:

```
class ShiftRegister extends Module {
  val io = IO(new Bundle {
    val in  = Input(UInt(1.W))
    val out = Output(UInt(1.W))
  })
  val r0 = RegNext(io.in)
  val r1 = RegNext(r0)
  val r2 = RegNext(r1)
  val r3 = RegNext(r2)
  io.out := r3
}
```

Listing 2. 4-bit Shift-register in Chisel

Of course, there are plenty of details to unpack here, such as the casting to `W` (bit width) type, wiring implicated from using `RegNext` specifically and so forth, but these are beyond the scope of this paper. For further instructions on Chisel, the reader is referred to [10], [11]. In [10], Chisel and dependencies are installed to a local machine and a set of examples, problems and solutions is provided, while the basics are explained through an online, github-wiki-based tutorial. In [11] Chisel is presented through an interactive tutorial based on a jupyter notebook with assignments and instructions intervened. There's an online version of the bootcamp, as well.

Chisel, rather, its toolchain, is capable of emitting Verilog for three different backends to generate either: (a) a software simulator that can be fed binaries, (b) a bitstream for an FPGA, or (c) netlist to be further used by automatized toolchain to yield GDS mask patterns ready for ASIC production. The actual advantages of Chisel over today's HDLs are shown in the next section, where processor cores used in our course are presented.

V. PROCESSOR CORES

Students get to learn about computer architecture and VLSI design by building one simple core, then applying that knowledge and expanding on it by investigating and modifying two more cores, each of these with several flavors.

A. Hack

Hack is the name of a computer and that computers processor - both developed within the now famous *Nand2Tetris* course, taught by Shimon Shocken and Noam Nisan. The course covers all abstraction levels starting from designing basic combinatorial modules, starting from NOT and XOR gates, collecting them into subsystems such as memories and aritmetic logic units (ALU), all the way through writing assembly code, operating system and even designing a complete java-like language, writing its compiler and then using the whole pyramid to design and play games. All tools required for the course are available free of charge at the course webpage [12], as well as the companion book [13], and the course itself, having been taught at several famous universities, is available at coursera.org.

We use the first half of Nand2Tetris, where students through six project assignments develop the chipset first, then the ALU, registers, RAM and program counter and learn about assembly code. Finally, they put all these together to obtain a fully functional Harvard architecture-based computer. In this way, a very sound foundation is created, making each candidate ready for tackling more complex problems such as memory-mapped input/output (MMIO) peripheries, pipeline, multicore processors, SoC design, etc.

B. Sodor

Sodor [14] is a collection of five simple and open-source cores written in Chisel, developed and published to be used at university level courses, as the basis for practical examples based on the theory presented in [3] and taught at UC Berkeley, *CS152/252A* courses [15].

After having developed Hack completely using the HDL provided at [12], and then taking a week to learn Chisel basics via the bootcamp [11], students learn about a real, RISC-V compatible processor. We focus on 1- and 5-stage implementations, as these provide enough to learn about performance metrics, most important benchmarks, measuring and comparing thus yielded numbers and introduce pipeline.

C. Rocket-chip

All the cores available in Sodor collection are written in ”plain” Chisel, meaning that no advanced principles such as core parameterization via diplomatic design patterns [16] are utilized to fully leverage advantages of Chisel over the standard HDLs.

Rocket-chip [17] is an open-source SoC generator also developed at the UC Berkeley, with the difference to Sodor cores that hardware it generates has actually been fabricated in silicon; therefore, while useful for research and teaching, these designs are applicable in industry, as well. As with other Chisel designs, its output is synthesizable Verilog and in this codebase the main advantages of Chisel are demonstrated.

It it important to differentiate between the *rocket* core as one of the cores that may be used within the SoC generated by the *Rocekt-chip* generator, whereas different cores may be used as well. Furthermore, the actual capabilities and performance may be fine tuned through the configuration parameters. This level of flexibility is yielded by the system’s modular design, built upon HCL approach to hardware design.

To learn about specifics of the following example in detail, we refer the reader to Chipyard documentation [18], as such discussion is beyond the scope of this paper. Here, we list a few lines of code in order to demonstrate the agility of the SoC generator, Chisel and the approach in general. The default configuration of a Rocket-chip will yield a single 64-bit core accompanied by a floating point unit (FPU) and with level 1 cache. However, if it is required to generate core with smaller cache and without the FPU (lower the energy consumption, decrease area), by inspecting the `Configs.scala` file an appropriate configuration may be found:

```
class DefaultSmallConfig extends
    Config(new WithNSmallCores(1) ++
    new WithCoherentBusTopology ++
    new BaseConfig)
```

Listing 3. Configuration to gneerate singlecore system without FPU and with smaller cache

Next, just by running:

```
make CONFIG=DefaultSmallConfig
```

we obtain synthesizable Verilog and a software simulator, with characteristics defined by the configuration line above. Similarly, configurations with dual (or more) cores may be used, at 32- or 128-bit widths, etc.

We do not delve into diplomatic patterns that enable such level of modularity, but rather follow examples from UC Berkeley [15], where the students in their first computer architecture/VLSI course develop understanding of the mutual dependency between hardware design and its application, i.e. software that it executes. Hence the projects for this part of the course are related to the features of the pipeline and cache: we ask for performance measurement while a specific piece of code is executed, then we seek ways to optimize hardware design by, say, changing cache associativity, and, finally, the benchmarks obtained after the modifications are compared - number of misses and instruction per cycle (IPC), in this particular example.

VI. Automated Toolchain

Once the design is settled from the computer architecture point of view, i.e. either we’ve reached the requirements or the time is simply up, it is time to look for ways to materialize the idea in a real circuit. While FPGA is a valid destination for Chisel code, this is not the topic of VLSI courses in general, so we focus on the ASIC digital synthesis toolchain within this paper. That is a set of software tools used to transform the Verilog (emitted by Chisel in this case) netlist into a physical digital circuit. In semiconductor industry today, these tools are vast proprietary suites developed and licensed by large companies such as Cadence or Synopsis. These are expensive to the point that quite a limited number of long running IC manufacturers may obtain the licenses on a regular basis. There are university and start-up programs, but those are also

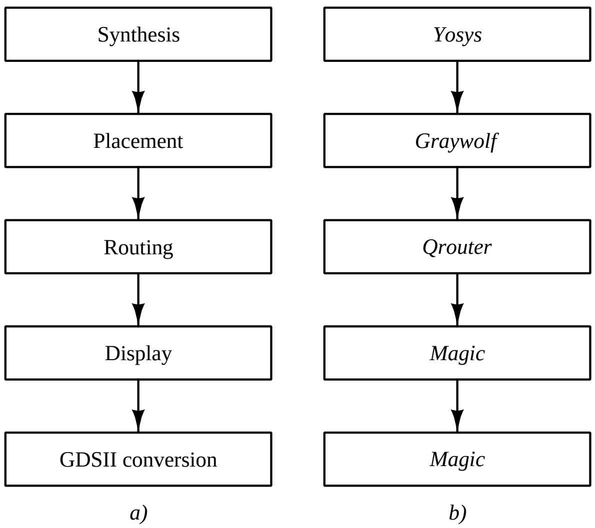


Fig. 1. Digital IC design open source toolchain: a) general approach, and b) tools applied during this course

far from free of charge for small universities or a team of two just starting out.

Qflow [19] is a complete, free of charge and open-source toolchain for synthesizing digital circuits starting from Verilog source and ending in physical layout for a specific target fabrication process. While the process is more detailed, we present a simplification describing only the major steps - rougly shown in Fig. 1. The first step in the automation process is to map the netlist onto a standard cells library, colloquially referred to as *PDK* (stemming from *project design kit*). PDKs are a topic of its own and a complex one, while at that, since these are also proprietary in general. Recently, there have been revolutionary development with SkyWater PDK [20], but not in time to be included in the course edition we are reporting on with this paper. In previous iterations of the course scalable CMOS PDK [21] was used. For now, we keep to the open-source PDK provided by the Oklahoma State University (OSU). This step is done by yosys [22]. Next, the design is to be placed and routed. In shortest terms, this when the standard cells are spread across the available area, while grouped in blocks and interconnected (routed). Graywolf [23] is the member of the qflow toolchain that does the placement, while routing is performed by qrouter [24]. Finally, for layout inspection, DRC and GDS generation Magic [25] is used. Qflow, nor the provided PDK are not capable of creating a microprocessor that may compete with current 3 GHz+ multicore server processors, but these tools will successfully handle simpler designs that may be found in SoC all over the market - such as SPI, for example. Firt live demonstration of a chip fabricated using nothing but qflow is presented in [26].

VII. Conclusion

While ISAs and processor cores are not directly a subset of a VLSI related course, we do live at a revolutionary moment in technology history - free and open source tools and PDKs are a reality, hence ASIC design and fabrication are within reach to individuals, start-ups and small universities with very limited funds. In future iterations of these courses, we plan to improve our automated design flow replacing qflow with openLane and including the open-source RAM compiler, openRAM. To overcome the steep learning curve of the new hardware design paradigm introduced with Chisel and rocket-chip, we are envisioning a Chisel GUI. Finally, we hope to fabricate students' designs through SkyWater 130 nm process, an open-source PDK.

References

[1] W. Shockley, "The theory of p-n junctions in semiconductors and p-n junction transistors," *Bell System Technical Journal. Institute of Electrical and Electronics Engineers (IEEE)*, vol. 28, no. 3, pp. 435–489, 1948.

[2] G. E. Moore, "Cramming more components onto integrated circuits," *Electronics*, vol. 38, no. 8, 1965.

[3] J. L. Hennessy and D. A. Patterson, *Computer architecture: a quantitative approach, 6th*. Morgan Kaufmann Publishers Inc., 2019.

[4] A. Pajkanovic, "On the application of free CAD software to electronic circuit curricula," in *Proc. of 3rd International Conference on Electrical, Electronic and Computing Engineering IcETRAN 2016*, Zlatibor, Serbia, 2016, pp. ELI1.3.1–4.

[5] A. Pajkanovic and Z. Ivanovic, "A report on recent development in application of free CAD software to IC curricula," in *Proc. of 5th IcETRAN 2018*, Palic, Serbia, 2018, pp. 847–851.

[6] A. Pajkanovic, "CMOS IC from schematic level to silicon within IC curricula using free CAD software," in *Proc. of INDEL 2020*, Banja Luka, Bosnia and Herzegovina, 2020.

[7] A. Waterman, "Design of the RISC-V instruction set architecture," Ph.D. dissertation, University of California at Berkeley, 2016. [Online]. Available: https://people.eecs.berkeley.edu/ krste/papers/EECS-2016-1.pdf

[8] A. Waterman, Y. Lee, D. A. Patterson, and K. Asanović, "The risc-v instruction set manual, vols. i-ii," EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2016-118, May 2016.

[9] J. Bachrach, H. Vo, B. Richards, Y. Lee, A. Waterman, R. Avižienis, J. Wawrzynek, and K. Asanović, "Chisel: Constructing hardware in a scala embedded language," in *DAC2012*, San Francisco, USA, 2012.

[10] chisel-tutorial. [Online]. Available: https://github.com/ucb-bar/chisel-tutorial/wiki

[11] chisel-bootcamp. [Online]. Available: https://github.com/freechipsproject/chisel-bootcamp

[12] nand2tetris. [Online]. Available: nand2tetris.org

[13] N. Nisan and S. Shocken, *The Elemenets of Computing Systems*. MIT Press, 2008.

[14] riscv-sodor. [Online]. Available: https://github.com/ucb-bar/riscv-sodor

[15] Computer Architecture 152/252A, spring 2021. [Online]. Available: https://inst.eecs.berkeley.edu/ cs152/sp21/

[16] H. Cook, "Diplomatic design patterns: A TileLink case study," in *CARRV17*.

[17] K. Asanović, R. Avižienis, J. Bachrach, S. Beamer, D. Biancolin, C. Celio, H. Cook, P. Dabbelt, J. Hauser, A. Izraelevitz, S. Karandikar, B. Keller, D. Kim, J. Koenig, Y. Lee, E. Love, M. Maas, A. Magyar, H. Mao, M. Moreto, A. Ou, D. Patterson, B. Richards, C. Schmidt, S. Twigg, H. Vo, , and A. Waterman, "The rocket chip generator," EECS Department, University of California, Berkeley, MA, Tech. Rep. Technical Report UCB/EECS-2016-17, 2016.

[18] Chipyard. [Online]. Available: https://chipyard.readthedocs.io/

[19] Qflow. [Online]. Available: http://opencircuitdesign.com/qflow/

[20] SkyWater. [Online]. Available: https://www.skywatertechnology.com/

[21] MOSIS. [Online]. Available: https://www.mosis.com/

[22] C. Wolf, "Yosys open synthesis suite," http://www.clifford.at/yosys/.

[23] "Graywolf," https://github.com/rubund/graywolf.

[24] "Qrouter," http://opencircuitdesign.com/qrouter.

[25] Magic. [Online]. Available: http://opencircuitdesign.com/magic

[26] T. Edwards and M. Kassem, "The Raven Chip: First-time silicon success with qflow and efabless," in *Free Silicon Conference 2019, FSiC 2019*, Paris, France, 2019. [Online]. Available: https://wiki.f-si.org/index.php/FSiC2019

Two approaches to automatic configuration of RS-485 network

Nikola Cvetković, Pavle Milenković, Nenad Jovičić and Vladimir Rajović

***Abstract*—The purpose of this paper is to provide an overview of common approaches to automatic configuration of half-duplex RS-485 network, as well as to introduce two alternative methods of automatic slave address configuration in a network. Described mechanisms will be analyzed and compared in terms of hardware and software complexity, while taking into consideration system robustness and implementation feasibility.**

***Index Terms*—RS-485, automatic configuration, differential bus, communication network**

I. INTRODUCTION

The interface commonly known as RS-485 represents an electrical standard used in serial communication systems. It only defines electrical characteristics of drivers and receivers connected into a network [1], thus leaving the opportunity of using various standardized or user-defined data communication protocols. Some of the examples of frequently used protocols include Modbus, used in industrial settings, and BACnet, often applied in automation of buildings and other monitoring applications [2]. RS-485 is a balanced differential electrical bus, either full-duplex or half-duplex, supporting up to 32 unit loads, where each unit load represents an impedance of approximately 12 kΩ. Fig. 1 shows an example of a common balanced system. It consists of a driver D and a receiver R. A termination resistor R_T is used on the input of the receiver, to match the input impedance of the lines.

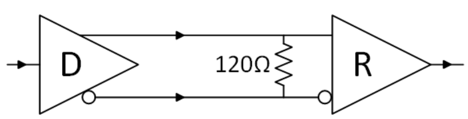


Fig. 1. A common balanced system

Since the lines of the bus are balanced, meaning they have equal impedances along their length and equal impedances with respect to ground, noise induced in the conductors and their electromagnetic radiation is minimized. Drivers' minimal differential output voltage has to be greater than 1.5 V across a load resistance of 54 Ω, while the minimal detectable differential input voltage of receivers is 200 mV. Described characteristics allow for a conventional transmission speed of up to 10 Mb/s and the maximum range of 1200 meters. When designing a system, an empirical relation for transmission speed and bus length ratio has to be taken into consideration [3]. Although the standard recommends 10 Mb/s, today's fast interface circuits are optimized for data rates of up to 50 Mb/s [4].

All the approaches for the automatic network configuration that will be analysed in this paper use a master/slave model of communication, in which only the master can initiate communication with the slave. The master can either send a broadcast address in order to address all the nodes in the network, or it can address individual slave by sending a specific address.

RS-485 standard supports various network topologies, the simplest of which is point-to-point connection (Fig. 2).

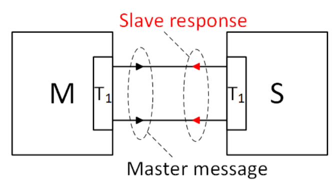


Fig. 2. Point-to-point topology

Adding additional network nodes to the topology mentioned above requires insertion of junction boxes. This modification, referred to as Backbone with Stubs, consists of a main differential bus called a backbone, branching off to the slaves with stubs (Fig. 3). The described topology is one of the two most commonly used methods for connection of the nodes.

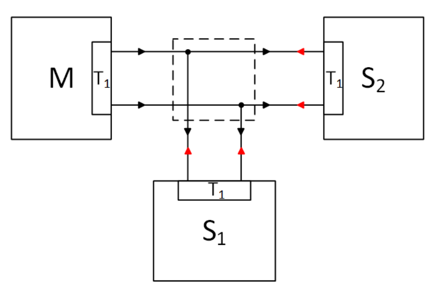


Fig. 3. Backbone with stubs topology

The second most frequently used connection method is daisy-chain topology (Fig. 4), which requires each two consec-

Nikola Cvetković is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: cn203094m@student.etf.bg.ac.rs).

Pavle Mileknović is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: mp203016mm@student.etf.bg.ac.rs).

Nenad Jovičić is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: nenad@etf.bg.ac.rs).

Vladimir Rajović is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: rajo@etf.bg.ac.rs).

utive nodes to be directly connected to each other via dedicated connection ports.

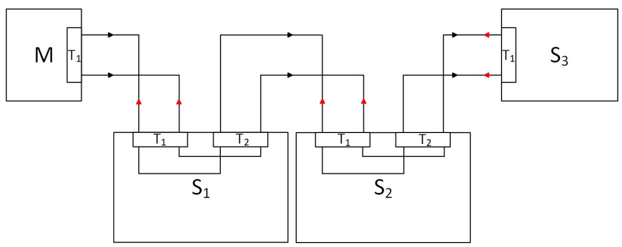


Fig. 4. Daisy-chain topology

When designing a network, special care has to be taken in order to provide a unique address for each node. In conventional address assigning methods, problems occur in manual work cost, hardware and software complexity, and inconvenience in network modification, including faulty node removal and network expansion. Without loss of generality, the main focus of the analysis will be on half-duplex RS-485 bus. Additionally, communication is performed in a traditional master-slave manner and information about relative physical position of each node has to be available at all times.

II. Common Slave Addressing Approaches

In practice, there are three widely used methods for node address configuration, which will be covered in more detail in the following subsections. These include different hardware address assignments, firmware hard coding of an address and software address assignment upon addition of a new node, one by one. The common characteristic of all these methods is that they are all in need of node configuration prior to address assignment.

A. Hardware Address Assignment

The simplest way to assign an address to a slave from a software perspective is to provide the hardware for address configuration. This can be achieved by using dual in-line package (DIP) switches, rotary switches, on-board jumpers, or by directly connecting every slave's address pin to either high or low logic levels. A typical configuration using a DIP switch is represented by Fig. 5.

In spite of being easy to implement, relatively cheap to design and in some applications easily reconfigurable, this method has considerable setbacks. This system is prone to human error, since all switches/jumpers rely on manual configuration, increasing the possibility of multiple devices having the same address. Additionally, in an environment where network nodes are not easily accessible, modification of the addresses might pose an issue. In some cases, major flaw of described approach can be the price of manual work required for network initialization and modification. Despite having the mentioned limitations, this approach is frequently found in commercial applications [5], [6].

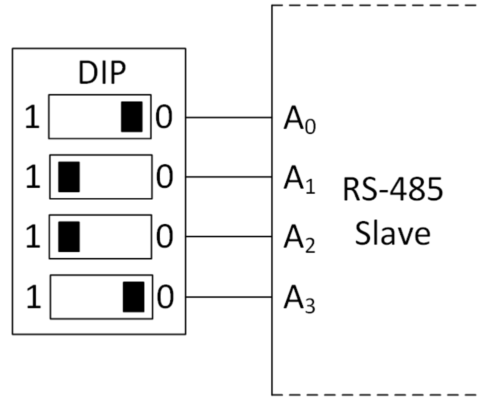


Fig. 5. Address configuration using a DIP switch

B. Firmware Hard Coding of an Address

Another method is the hard coding of an address inside the slave firmware. This approach results in a simpler and cheaper design, as it does not require any additional hardware. Since it does not require a direct human contact during deployment in order to configure the slave, the system is less prone to human error than the previous approach. However, the main weakness of this method lies in the effort to write and maintain different versions of firmware for each node in the network. In terms of network extension and modification, nothing is improved compared to the hardware assignment method.

C. One by One Network Expansion

If hardware modification of the nodes is not possible, different firmware versioning can be avoided by gradual expansion of the network, one node at the time. A common algorithm of RS-485 bus initialization is represented by Fig. 6.

At the beginning of the initialization of the network, every slave is configured to have the same default address, e.g. address 0x00, while the initial topology resembles point-to-point communication (Fig. 2). Master sends a broadcast message, addressing all slaves with the default address. Since the nodes are being added to the network one by one, at any given moment, there will be no more than one device with the default address. The node with the default address sends an echo message, requesting the new address from the master. Finally, the master sends a message containing the new address of the slave and waits for the confirmation of the message, checking if the slave is configured properly. Master can be configured to send a broadcast message periodically, or by user command upon connection of a new node. Some implementation of a waiting/timeout loop should be considered, since various errors in a network can result in improper signal reception, either on master or slave side of the connection. Another advantage of this method is the simplicity of adding new nodes into a network - no additional work to adjust the software or configure the switches is required, apart from physically connecting the nodes one after another. Although relatively simple in terms of hardware and software complexity from the

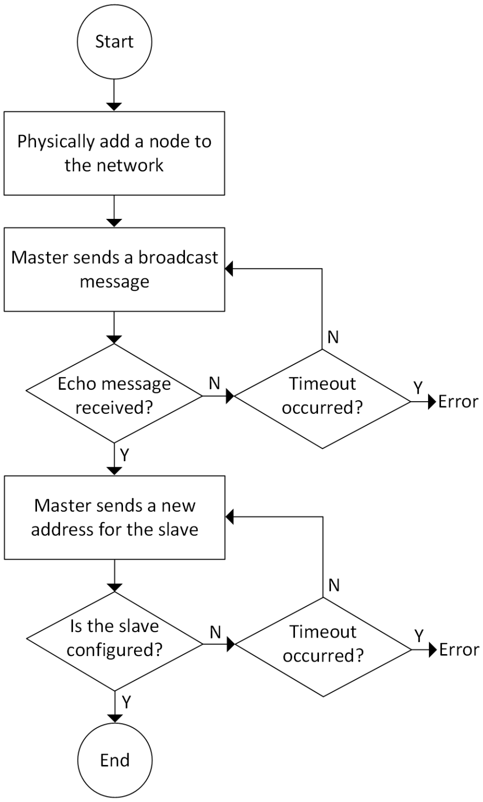


Fig. 6. One by one initialization algorithm

slave side, this approach requires a long initialization process, which results in a costly manual work.

III. Proposed Solutions

In the previous sections, the internal structure of network nodes was implemented using a single RS-485 transceiver. A typical example of an RS-485 transceiver can be observed in the Fig. 7.

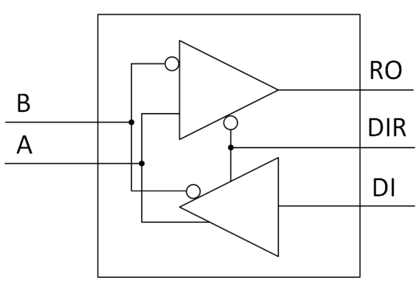


Fig. 7. Typical RS-485 transceiver

By designing the network nodes using additional RS-485 transceiver, new methods for automatic address assignment can be formulated.

A. Utilizing Network Topology for Slave Addressing (Domino Method)

Proposed structure of a network node is given in the Fig. 8. This solution offers a flow of information through the slave, by

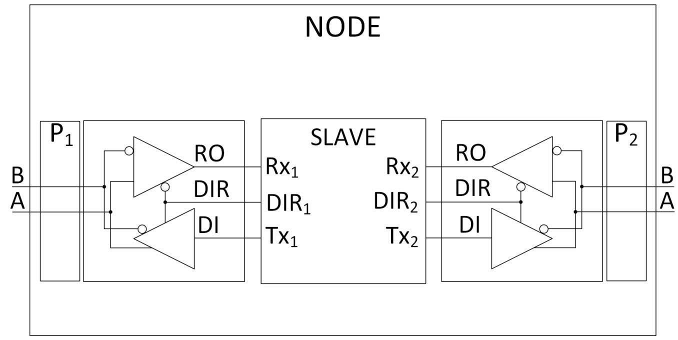


Fig. 8. Node structure for the Domino method

making use of two serial ports. The line topology obtained by connecting multiple slaves with previously presented structure is shown in Fig. 9. Nodes are denoted by N_i , $i = \overline{0, n-1}$, while each node has two ports, P_1 and P_2.

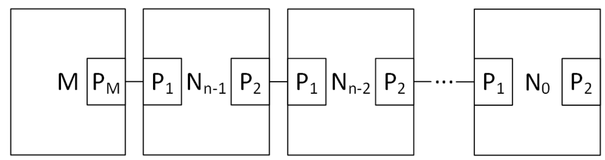


Fig. 9. A proposed network topology for the Domino method

The basic principle of the proposed solution is that there is a predetermined address that triggers the responses from every slave, similarly to the broadcast address in any multiple-access communication network. In a given example, the value of predetermined address is 0x00. The system initialization phase starts with determining the number of nodes in the network. Depending on the expected number of nodes, either incremental or decremental approach can be used. Perhaps more intuitive is the incremental approach, used for small-scale networks. The algorithm starts by master sending a test message over P_M port, containing the predetermined address, which stores the information about nodes counted thus far, in this case 0x00. Since the predetermined address of the slave matches the address field in the message, the first slave in the line responds by echoing the message back to the master, over port P_1. The master detects an echo, and calculates the total number of detected devices by incrementing an address field by one. Now the master repeats sending a test message, this time with the address field 0x01. The first node in the line will not respond to the test message, since its predetermined address is 0x00. Because there is no match, the node will decrement only the address field, and pass the message to the next node in the line through the P_2 port. After passing the message, the node will change the direction of information flow, by toggling the values of DIR_1 and DIR_2 pins, thus configuring the P_1 port as a driver and P_2 port as receiver. The second node receives a message on P_1 port, now with the address 0x00, and responds to it by sending the echo, using the same port and address 0x00. The preceding node receives the echo on its P_2 port, increments the address field to 0x01, and passes the message to the master using P_1 port. Upon receiving an echo, the master increments the address field, updating the number of counted nodes to 0x02. If the total number of nodes in the line is given by n, the described process repeats for a total of n times. Successful addressing of the n-th node is presented in Fig. 10.

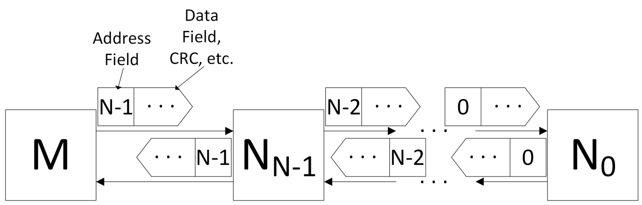


Fig. 10. Domino method addressing

Lastly, the master sends a test message once again, this time reaching the last node in the line, with the address of 0x01. The last node receives the message with this address field, decrements it and attempts to send it using its P_2 port to the next node. Since there are no nodes left, there will also be no echo message back to the master. The master waits for the echo for a certain amount of time, after which a timeout event occurs, signaling the master that there are no nodes left.

Once the total number of devices is determined, it can be used for communicating with each device. This communication protocol is based on master sending a message with the address field corresponding to the relative position of the node in the line.

One of the main issues with the described approach lies in addition of new nodes into a network. When the node is added, the initialization routine has to be repeated. This may be performed by the request from the user, or by periodical repetition of the algorithm. The suggested approach would be to manually start the routine, as it is less time-consuming.

The other downside of the proposed solution is the handling of a malfunctioned node or electrically corrupted bus - the line will be interrupted as the signal can not propagate through subsequent nodes.

Domino method occupies two serial ports of the node's micro-controller, which can be a limiting factor in system design, potentially increasing the overall cost of the system. Additionally, this solution introduces a significant computational delay, as the information is processed by each micro-controller it passes through. The total delay added into a single communication cycle, consisting of a single master request and slave response when addressing i-th slave in the line, is given by:

$$T_{DELAY_I} = 2 \cdot i \cdot T_{MCU}, \tag{1}$$

where T_{MCU} is the time required for a message to be processed and/or modified. This value has to be multiplied by the number of nodes that information passes through, and, since the message propagates back to the master, the whole value is multiplied by two. When evaluating T_{MCU}, it is important to include not only the time required for increment/decrement operations, but also time required for receiving and sending a modified message using a serial port. Assuming there are n nodes in the line, the maximum delay introduced by the Domino method is given by the expression:

$$T_{MAX} = 2 \cdot n \cdot T_{MCU}. \tag{2}$$

B. Staged Address Assignment with Bus Bridging (Pontoon Method)

As it can be observed from the previously described method, the core of the node constantly processes received information, introducing delay. This potential overload of the slave's micro-controller can be avoided by bridging the RS-485 bus inside a node using a discrete multiplexer (Fig. 11).

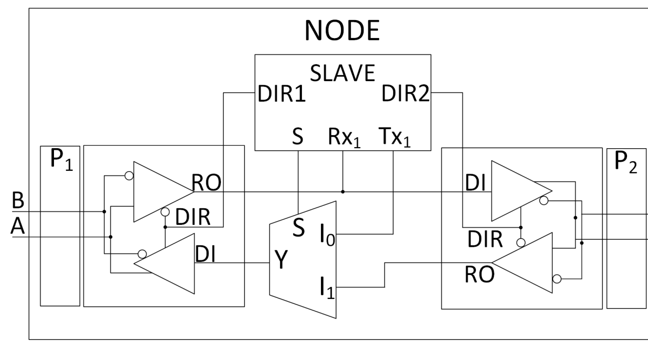


Fig. 11. Node structure for the Pontoon method

On a system level, the topology is identical to the one described in Fig. 9. Contrary to the previous approach, addresses of the nodes can be assigned by the master and each address has to be stored inside of the corresponding node. Additionally, only one serial port is used, which allows for more flexibility when designing a system. Initially, every node of the network is configured in a way that both of its ports receive the information, and each node's address is set to a predetermined value, e.g. 0x00. The master starts the process of address assignment by sending a message, addressing a device with the predetermined address. The first node receives the message, recognizing its address. Since its P_2 port is in receive mode, the propagation of the message stops. After recognizing its address, the first node sends an echo to the master by selecting I_0 multiplexer input and changing the direction of P_1 port, asking for a new address. The master sends another message for slave address configuration, where the address field can be arbitrarily determined by the user. Upon receiving a new address, the first node sends back a conformation message to the master and changes the direction of P_2 port in order to allow for the next address configuration message to pass through it. The master acknowledges that the first node is successfully configured and sends the message with the address 0x00 once again. The first node receives this message on P_1 port, and since its P_2 port is configured as a driver, the message can pass through the node, without unnecessary computing delay. As the information is passed, both ports of the first node have to change directions and I_1 multiplexer input has to be selected, to allow for opposite data flow. The second node in the line receives the message on its P_1 port. Because its address is 0x00 and its P_2 port is blocking the information flow, it echoes the message back to the master, requesting a new address. The first node receives an echo on its P_2 port, passes it through P_1 port and changes the direction of the ports once again. The process continues until the last node is reached, and once again, the master has to implement a waiting mechanism, when the last node tries to pass the configuration message. If a certain slave recognizes its address after the initialization is done, it has to configure its P_1 port as a driver, selecting the I_0 multiplexer input.

This time, adding new nodes to the network does not require repetition of the whole initialization process, but only a single passing of the configuration message. Again, the line will be interrupted in a situation where the node is malfunctioning or the bus is somehow corrupted, either by open or short circuit.

Contrary to the Domino method, this approach introduces additional delay during network configuration. For any node in a network, configurational delay is:

$$T_{CNFG_{SINGLE}} = 2 \cdot T_{MCU}, \tag{3}$$

where, again, T_{MCU} is the time required for a message to be received, processed and transmitted. It takes single T_{MCU} to process a message containing the predetermined address, and another T_{MCU} to process a message containing a new node address. For a network containing n nodes, the total configurational delay is given by:

$$T_{CNFG_{TOTAL}} = 2 \cdot n \cdot T_{MCU}. \tag{4}$$

After the network is successfully configured, the duration of a communication cycle, consisting of a message transmission from the master and a response from any slave is:

$$T_{COMM} = T_{MCU}, \tag{5}$$

since the message passes through preceding nodes directly, without being processed by their micro-controllers.

IV. Conclusion

From the hardware perspective, both automatic configuration methods presented require an additional RS-485 transceiver, which increases design complexity and overall price of each node. However, these approaches completely eliminate the need for manual work, except when adding or removing a node. By using the described algorithms for node communication, the whole network can be effortlessly reconfigured through master. This means that the nodes do not have to be easily reachable, which avoids implementing galvanic isolation of RS-485 transceivers, except when common mode voltage reduction is required. Further steps can be taken in order to completely relinquish the use of galvanic isolation [7].

The Domino approach that utilizes the network topology for slave addressing does not require individual storing of a slave's address. Therefore, it can be used in a system where there are multiple slaves of the same type, often performing a similar task. The example of a system that can benefit from this solution can be temperature monitoring networks, or street light management systems, where each node is functionally the same.

Additionally, this solution is less complex hardware-wise than the Pontoon method, but occupies an additional serial port and introduces computational delay, as the micro-controller processes address fields of the message.

Pontoon approach allows each node to have a unique address, independently of its position in a communication line. This can be useful in a system where certain types of sensors or devices have to share a specific address range, for example in various industrial or home automation systems.

Since there is no need for master to periodically poll the network in order to detect new devices, the whole method of network reconfiguration, adding and removing the nodes, is simplified compared to the first proposed method. This approach introduces virtually no additional communication delay, apart from the propagation of a signal through multiplexer. Pontoon approach does introduce a multiplexer in the node's design, but offers a possibility of compromise price-wise, since it occupies a single serial port, allowing for a simpler micro-controller. The usage of a discreet multiplexer can be avoided by directly connecting RO line from P_2 port to the DI line from P_1 port and Tx_1 line, and by keeping Tx_1 line in a high-impedance state whenever the message is expected to be received from P_2 port.

The Domino method does not require node address configuration, meaning that there is no time delay introduced during network initialization, compared to the Pontoon method that introduces a delay of $2nT_{MCU}$ during this phase, for a network consisting of n nodes. On the other hand, during a communication cycle, the Domino method requires $2iT_{MCU}$, for i-th node in the line, while the Pontoon approach requires only T_{MCU} for the same operation, regardless of the node position.

Both of the proposed solutions can be used in systems that are inherently linearly connected. In addition to the examples mentioned before, those systems include metallurgical furnaces, speed and traffic detection, environmental, and other kinds of roadway sensors, as well as various sewer sensors. Furthermore, the described methods can relatively easily be generalized in order to be applied to other communication systems, using different electrical standards and protocols. Without change in principal node structure, approaches can be easily adjusted to fit other serial standards. These methods of address assignments can be adapted for different communication mediums, including wireless and optical communication.

The future work should cover defining methodology for generalization of the presented address assignment approaches, as well as finding ways to improve system robustness and failure recovery capabilities.

ACKNOWLEDGMENTS

This work was supported by the Ministry of Education, Science and Technological Development, Republic of Serbia, under Projects TR32043 and TR32039.

REFERENCES

[1] *Electrical Characteristics of Generators and Receivers for Use in Balanced Digital Multipoint Systems*, ANSI/TIA/EIA-485-A-1998, TIA/EIA, 1998.

[2] J. Axelson, *Serial Port Complete: COM Ports, USB Virtual COM Ports, and Ports for Embedded Systems*, 2^{nd} ed. Madison, WI, USA: Lakeview Research LLC, 2007.

[3] T. Kugelstadt, "Application Report - The RS-485 Design Guide," TI, Dallas, TX, USA, 2021.

[4] Texas Instruments, "SN65HVD7x 3.3-V Full-Duplex RS-485 Transceivers with ±12-kV IEC ESD," TI, Dallas, TX, USA, 2019.

[5] REFLEX Thinking Solutions, "Busmodul - Profibus DP, Operating Manual," Reflex Winkelmann GmbH, Ahlen, Germany, 2014.

[6] C. Boiano, P. Guazzoni, L. Zetta, C. Guazzoni, and A. Pagano, "A 16-channel programmable antialiasing amplifier", IEEE Nuclear Science Symposuim & Medical Imaging Conference, Knoxville, TN, USA, 30 Oct. - 6 Nov. 2010.

[7] N. S. Jovičić and V. M. Rajović, "A Floating Linear Voltage Regulator for Powering Large-Scale Differential Communication Networks," *IEEE Access*, vol. 6, pp. 24669-24679, May, 2018.

МЕТРОЛОГИЈА
/
METROLOGY
(МЛ/MLI)

Obezbeđenje validnosti rezultata ispitivanja nivoa snage smetnji ponavljanjem merenja

Aleksandar M. Kovačević, Nenad Munić, Veljko Nikolić, Ljubiša Tomić

Apstrakt—**Akreditovana laboratorija mora da preispituje svoje rezultate ispitivanja radi obezbeđenja njihove validnosti. Jedan od načina je ponavljanje ispitivanja na poznatom uređaju. U radu je prikazano merenje nivoa snage smetnji na kuvalu za vodu. Pri tome, praćenje validnosti rezultata ispitivanja se obavlja u skladu sa utvrđenom procedurom.**

***Ključne reči*—Validnost; merenje; snaga smetnji.**

I. Uvod

AKREDITOVANA laboratorija mora da preispituje svoje rezultate ispitivanja radi obezbeđenja njihove validnosti. Pri tome, ona mora da ima proceduru za praćenje validnosti rezultata [1].

Ovo praćenje mora da se planira i preispituje i mora da obuhvati sledeće [1]:

a) korišćenje referentnih materijala ili materijala za kontrolu kvaliteta;
b) korišćenje alternativnih instrumenata koji su etalonirani i istog nivoa tačnosti tako da daju sledljive rezultate;
c) funkcionalnu proveru merne opreme i opreme za ispitivanje;
d) korišćenje etalona za proveru ili radnih etalona sa kontrolnim kartama, tamo gde je to primenljivo;
e) međuprovere merne opreme;
f) ponavljanje ispitivanja ili etaloniranja korišćenjem istih ili različitih metoda;
g) ponovno ispitivanje ili etaloniranje predmeta koji se čuvaju;
h) korelacija rezultata za različite karakteristike predmeta;
i) preispitivanje rezultata o kojima se izveštava;
j) međulaboratorijska poređenja;
k) ispitivanje slepih uzoraka.

Pri tome, svaka laboratorija obuhvata ono što joj je odgovarajuće.

Odeljenje za elektromagnetsku kompatibilnost i uticaje okoline je akreditovano u oblasti ispitivanja elektromagnetske kompatibilnosti (u daljem tekstu Odeljenje za EMC i uticaje okoline) u okviru Centra za ispitivanje proizvoda, Tehnički opitni centar (TOC) iz Beograda [2]. Tako, Odeljenje za EMC i uticaje okoline u sklopu obezbeđenja validnosti rezultata ispitivanja redovno vrši sledeće aktivnosti: funkcionalnu proveru merne opreme i opreme za ispitivanje, međuprovere ključne merne opreme, ponavljanje ispitivanja korišćenjem istih metoda, ponovno ispitivanje predmeta koji se čuvaju, preispitivanje rezultata o kojima se izveštava, međulaboratorijsko poređenje. U tu svrhu, u TOC-u je izrađen dokument Uputstvo za obezbeđenje poverenja u kvalitet rezultata ispitivanja [3], a u skladu sa standardom SRPS ISO/IEC 17025:2017 [1]. Neke od navedenih aktivnosti (međulaboratorijsko poređenje, međuprovere ključne merne opreme) su i prezentovane na nacionalnim naučnim skupovima [4, 5].

Aleksandar M. Kovačević – Fakultet tehničkih nauka, Univerzitet u Kragujevcu, Svetog Save 65, 32000 Čačak, Srbija (e-mail: aleksandar.kovacevic@ftn.kg.ac.rs). Nenad Munić, Veljko Nikolić – Tehnički opitni centar, Generalštab Vojske Srbije, Vojvode Stepe 445, 11000 Beograd, Srbija (e–mail: nenadmunic@yahoo.com; veljkozmaj@yahoo.com). Ljubiša Tomić – Vojnotehnički institut, Ratka Resanovića 1, 11000 Beograd, Srbija (e-mail: ljubisa.tomic@gmail.com).

U ovom radu je prikazano ponavljanje merenja nivoa snage smetnji na kuvalu za vodu (aparat za domaćinstvo), koja su realizovana 2017. i 2018. godine, respektivno. Merenja nivoa snage smetnji su obavljena na mrežnom vodu (niskonaponska elektroenergetska mreža) prema standardu SRPS EN 55014-1:2010/A1:2010/A2:2012 [6, 7, 8] i SRPS EN 55014-1:2017 [9].

Cilj ponavljanja ispitivanja (merenje nivoa snage smetnji) na poznatom uređaju je da, na osnovu analize dobijenih rezultata i zadatog kriterijumima, Odeljenje za EMC i uticaje okoline obezbedi njihovu validnost. Naime, podaci dobijeni na osnovu ove analize se koriste za upravljanje i poboljšavanje aktivnosti Odeljenja za EMC i uticaje okoline. Ukoliko se nađe da su rezultati analiza dobijenih podataka izvan prethodno definisanih kriterijuma, mora da se preduzima odgovarajuća mera kako bi se sprečilo da se izveštava o netačnim rezultatima [1].

II. Uslovi ispitivanja

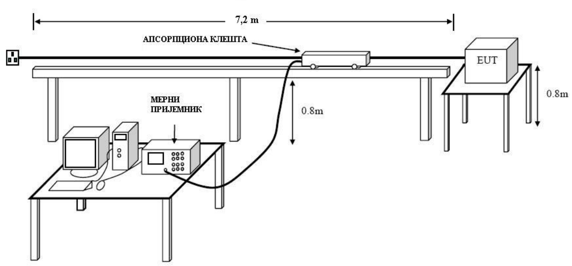
Merenje nivoa snage smetnji je vršeno na mrežnom vodu kuvala za vodu prema standardu [6, 7, 8, 9]. Tip kuvala za vodu je FA-5428-2, proizvođača „TZS FIRST Austria“, ser.br. SP-3711, 220 V, 50 Hz, 2200 W.

Kuvalo za vodu je postavljeno na ugao drvenog (neprovodnog) stola da bi apsorpciona klešta bila što bliže drvenoj klupi (Sl. 1). Pri tome, mrežni vod uređaja je bio priključen na pravolinijski deo mrežnog voda, kako bi se u svakom trenutku omogućilo pomeranje apsorpcionih klešta (Sl. 2).

Apsorpciona klešta su bila postavljena oko voda tako da je merena veličina srazmerna snazi smetnji na vodu [6, 7, 8, 9, 10]. Da bi se mogli porediti rezultati merenja, nije vršeno menjanje mernog položaja apsorpcionih klešta, tj. obeležen je položaj klešta u odnosu na pravolinijski deo mrežnog voda.

Sl. 1. Povezivanje kuvala za vodu na pravolinijski deo mrežnog voda.

Pre početka merenja snage smetnji izvršena je kontrola smetnji okoline kada je uređaj isključen (nivo smetnji ambijenta).



Sl. 2. Merenje snage smetnji na mrežnom vodu.

Merenje nivoa snage smetnji je vršeno na mrežnom vodu u frekvencijskom opsegu od 30 MHz do 300 MHz, na 3 izabrane frekvencije, iz 3 frekvencijska podopsega (30 MHz do 100 MHz, 100 MHz do 200 MHz, 200 MHz do 300 MHz), respektivno, kako je propisano standardima [6, 7, 8, 9, 10].

Na početku merenja se prvo vrši pretraživanje ili prebrisavanje celokupnog opsega (prvo merenje), za šta se koristi vršni – Pk detektor analizatora spektra [6, 7, 8, 9]. Zatim se za svaku navedenu frekvenciju (odnosno za definisani podopseg oko nje) izvrši drugo merenje sa sva tri detektora analizatora spektra (vršni – Pk detektor, kvazivršni – QP detektor, detektor srednje vrednosti – AVG detektor).

Pri tome, kuvalo za vodu je bilo u radu (grejalo je vodu).

Za navedena merenja korišćena su sledeća merna sredstva i oprema:

- EMC analizator spektra E7402A, „AGILENT“, od 100 Hz do 3 GHz,
- Apsorpciona klešta AMZ 41A, Teseq, od 30 MHz do 1 GHz,
- Stoni računar ASUS sa aplikacijom za automatizaciju merenja (EMC Measurement Application E7415A);
- BNC kabl, Teseq,
- Kabl RG-214/U.

Pri tome, karakteristike merne opreme zadovoljavaju propisane standarde [11, 12].

Uslovi okoline:

- temperatura okoline: 21 °C ± 2 °C,
- relativna vlažnost vazduha: 65 % ± 15 %.

III. Kriterijum za ocenu rezultata merenja

Za ocenu rezultata merenja korišćena je sledeća formula:

$$\left|X_2 - X_1\right| \le U_m \text{ – zadovoljavajući (prihvatljiv) rezultat} \quad (1)$$

gde su:

X_1 – izmereni nivo snage smetnji (dBpW), 2017. godine,

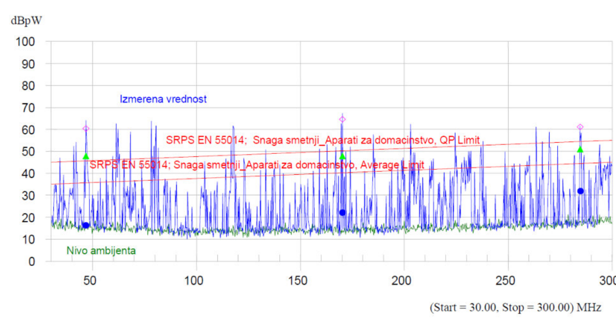
X_2 – izmereni nivo snage smetnji (dBpW), 2018. godine,

U_m – proširena merna nesigurnost merenja nivoa snage smetnji (za faktor proširenja ili prekrivanja $k = 2$).

Kao kriterijum za ocenu validnosti ponovljenih rezultata merenja uzeta je zadovoljenost (1). Naime, ukoliko je zadovoljenost (1) ispunjena, tj. razlika u dobijenim vrednostima je u granicama merne nesigurnosti, smatra se da je obezbeđena validnost rezultata merenja, pa nema potrebe za korektivnim merama. U suprotnom, potrebno je uvesti korektivne mere (npr. vanredno etaloniranje merne opreme i sl.).

IV. Rezultati merenja

Grafički prikaz rezultata merenja nivoa snage smetnji na mrežnom vodu (2018) je dat na Sl. 3.



Sl. 3. Grafički prikaz rezultata merenja nivoa snage smetnji na mrežnom vodu (2018).

Prvo je izvršeno pretraživanje ili prebrisavanje celokupnog opsega (prvo merenje) sa Pk detektorom analizatora spektra. Zatim se za svaku navedenu frekvenciju (odnosno za definisani podopseg oko nje) izvrši drugo merenje sa sva tri detektora analizatora spektra (Pk detektor, vrednosti na grafiku su označene rombićima; QP detektor, vrednosti na grafiku su označene trouglićima; AVG detektor, vrednosti na grafiku su označene kružićima).

Brojčane vrednosti su date u Tabeli 1 (2017) i Tabeli 2 (2018), respektivno.

TABELA I
REZULTATI MERENJA NIVOA SNAGE SMETNJI (2017)

f (MHz)	Detektor	X_1 (dBpW)
46	Pk	57,5
	QP	48,9
	AVG	17,0
170	Pk	66,4
	QP	48,6
	AVG	24,5
284	Pk	64,5
	QP	53,1
	AVG	33,7

TABELA II
REZULTATI MERENJA NIVOA SNAGE SMETNJI (2018)

f (MHz)	Detektor	X_2 (dBpW)
46	Pk	60,5
	QP	47,9
	AVG	15,9
170	Pk	64,4
	QP	47,6
	AVG	22,8
284	Pk	61,0
	QP	50,1
	AVG	30,7

Proširena merna nesigurnost, U_m, za merenje nivoa snage smetnji (k = 2) iznosi 4,183 dB (u radu je usvojena vrednost od 4,2 dB). Pri tome, obračun merne nesigurnosti za snagu smetnji dat je detaljno u internoj Proceduri za određivanje merne nesigurnosti kod ispitivanja EMC [13].

Na osnovu dobijenih rezultata (Tabela 1 i Tabela 2) i postavljenog kriterijuma formirana je Tabela 3.

TABELA III
VALIDNOST REZULTATA MERENJA NIVOA SNAGE SMETNJI

f (MHz)	Detektor	X_1 (dBpW)	X_2 (dBpW)	$\lvert X_2 - X_1 \rvert$ (dB)	U_m (dB)	$\lvert X_2 - X_1 \rvert \le U_m$
46	Pk	57,5	60,5	3,0	4,2	Da
	QP	48,9	47,9	1,0	4,2	Da
	AVG	17,0	15,9	1,1	4,2	Da
170	Pk	66,4	64,4	2,0	4,2	Da
	QP	48,6	47,6	1,0	4,2	Da
	AVG	24,5	22,8	1,7	4,2	Da
284	Pk	64,5	61,0	3,5	4,2	Da
	QP	53,1	50,1	3,0	4,2	Da
	AVG	33,7	30,7	3,0	4,2	Da

Iz Tabele 3 se vidi da je zadovoljenost (1) ispunjena.

Na osnovu svega, utvrđeno je da je razlika u dobijenim vrednostima u granicama merne nesigurnosti, tako da je obezbeđena validnost rezultata merenja, pa nema potrebe za korektivnim merama.

V. ZAKLJUČAK

Odeljenje za elektromagnetsku kompatibilnost i uticaje okoline iz Tehničkog opitnog centra iz Beograda, koje je akreditovano u oblasti ispitivanja elektromagnetske kompatibilnosti (EMC), svake godine vrši praćenje validnosti rezultata ispitivanja, koje se obavlja u skladu sa utvrđenom procedurom [3].

Ponavljanje ispitivanja na poznatom uređaju je jedan od postupaka za praćenje validnosti rezultata. Pri tome, cilj ponavljanja ispitivanja (merenje nivoa snage smetnji) na poznatom uređaju (kuvalo za vodu) je da, na osnovu analize dobijenih rezultata i zadatog kriterijumima, Odeljenje za EMC i uticaje okoline obezbedi njihovu validnost.

U ovom radu je prikazano ponavljanje merenja nivoa snage smetnji na kuvalu za vodu (aparat za domaćinstvo), koja su realizovana 2017. i 2018. godine, respektivno. Merenja nivoa snage smetnji su obavljena na mrežnom vodu (niskonaponska elektroenergetska mreža) prema standardima [6, 7, 8, 9]. Kao kriterijum za ocenu validnosti ponovljenih rezultata merenja uzeta je zadovoljenost (1), tj. da razlika u dobijenim vrednostima bude u granicama merne nesigurnosti.

Kako je utvrđeno da je razlika u dobijenim vrednostima u granicama merne nesigurnosti, to je obezbeđena validnost rezultata merenja, pa nema potrebe za korektivnim merama.

LITERATURA

[1] *Opšti zahtevi za kompetentnost laboratorija za ispitivanje i laboratorija za etaloniranje*, SRPS ISO/IEC 17025, ISS, 2017.
[2] http://www.toc.vs.rs.
[3] *Uputstvo za obezbeđenje poverenja u kvalitet rezultata ispitivanja*, Interni dokument, TOC, 2005.
[4] A. Kovačević, N. Munić, V. Nikolić, Lj. Tomić, I. Kostić, "Međulaboratorijsko poređenje merenja snage smetnji na mrežnom vodu", Zbornik radova 62. konferencije ETRAN-a, Palić, Srbija, str. 256-259, 11-14.06.2018.
[5] A. Kovačević, N. Munić, V. Nikolić, Lj. Tomić, I. Kostić, "Međuprovera EMC analizatora spektra između dva etaloniranja", Zbornik radova 63. konferencije ETRAN-a, Srebrno jezero, Srbija, str. 523-526, 03-06.06.2019.
[6] *Elektromagnetska kompatibilnost – Zahtevi za aparate za domaćinstvo, električne alate i slične uređaje – Deo 1 : Emisija,* SRPS EN 55014-1, ISS, 2010.
[7] *Elektromagnetska kompatibilnost – Zahtevi za aparate za domaćinstvo, električne alate i slične uređaje – Deo 1 : Emisija – Izmena 1,* SRPS EN 55014-1/A1, ISS, 2010.
[8] *Elektromagnetska kompatibilnost – Zahtevi za aparate za domaćinstvo, električne alate i slične uređaje – Deo 1 : Emisija – Izmena 2,* SRPS EN 55014-1/A2, ISS, 2012.
[9] *Elektromagnetska kompatibilnost – Zahtevi za aparate za domaćinstvo, električne alate i slične uređaje – Deo 1 : Emisija,* SRPS EN 55014-1, ISS, 2017.
[10] *Specifikacija aparata i metoda za merenje radio-smetnji i imunosti — Deo 2-2: Metode merenja smetnji i imunosti — Merenje snage smetnje*, SRPS EN 55016-2-2, ISS, 2012.
[11] *Specifikacija aparata i metoda za merenje radio-smetnji i imunosti – Deo 1-1: Aparati za merenje radio-smetnji i imunosti – Merni aparati*, SRPS EN 55016-1-1:2011/A1:2012/A2:2015, ISS.
[12] *Specifikacija aparata i metoda za merenje radio-smetnji i imunosti – Deo 1-3: Aparati za merenje radio-smetnji i imunosti – Pomoćna oprema – Snaga smetnji*, SRPS EN 55016-1-3:2011/A1:2017, ISS.
[13] *Procedura za određivanje merne nesigurnosti kod ispitivanja EMC*, Interni dokument, TOC, 2018.

Ensuring the validity of test results of disturbance power levels by repeating the measurement

Aleksandar M. Kovačević, Nenad Munić,
Veljko Nikolić, Ljubiša Tomić

ABSTRACT

An accredited laboratory shall review its test results to ensure their validity. One way is to repeat the test on a known device. The paper presents the measurement of disturbance power levels on an electric kettle. In addition, the monitoring of the validity of test results is performed in accordance to a defined procedure.

Očitavanje pseudoslučajnog koda pomoću linearnog niza fotodetektora kod pseudoslučajnih pozicionih enkodera

Ivana Ranđelović, Dragan Denić i Goran Miljković

***Apstrakt*— U radu se predlaže primena linearnog niza fotodetektora za paralelno očitavanje pseudoslučajnog binarnog koda kod apsolutnih pseudoslučajnih pozicionih enkodera. Predloženom metodom očitavanja koda bi se eliminisala potreba za inicijalnim kretanjem osovine enkodera prilikom njegovog startovanja u slučaju serijskog očitavanja koda. Urađena je simulacija rada predloženog rešenja pozicionog enkodera primenom softverskog paketa LabVIEW. Digitalna simulacija je realizovana u formi dva programa, pri čemu jedan program simulira sistem za očitavanje koda pseudoslučajnog pozicionog enkodera, dok drugi program simulira funkcionisanje elektronskog bloka takvog enkodera. Predstavljena je analiza primene komercijalno dostupnog linearnog niza fotodetektora za očitavanje koda sa realizovanog staklenog diska enkodera na kome je pseudoslučajna kodna traka.**

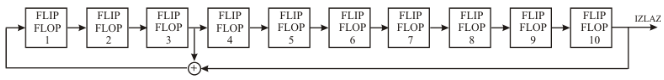
***Ključne reči*— pseudoslučajni pozicioni enkoder, merenje pozicije, paralelno očitavanje pseudoslučajnog koda, linearni niz fotodetektora, virtuelna instrumentacija, LabVIEW.**

I. UVOD

Za određivanje pozicije pokretnih sistema u industriji sve više se koriste pseudoslučajni pozicioni enkoderi koji predstavljaju značajan pravac u razvoju novih tipova apsolutnih enkodera. Enkoderi sa pseudoslučajnim kodom omogućuju visoku tačnost i dobru pouzdanost. Prilikom razvoja pseudoslučajnih enkodera potrebno je podjednako dobro realizovati sve njegove sastavne komponente: realizacija kodnog diska shodno optičkom čitaču koji će se koristiti, način očitavanja pseudoslučajnog koda [1], metoda skeniranja koda, konverzija pseudoslučajnog u prirodni kod [2], implementacija detekcije grešaka očitavanja koda [3], kao i postupak određivanja nulte pozicije prilikom montaže enkodera na osovinu motora [4].

Primena pseudoslučajnih pozicionih enkodera je široka. Mogu se naći u raznim oblastima: za merenje pozicije u industriji i robotici, pozicioniranje kranova i dizalica [5], za kontrolu kretanja automatski vođenih mašina [6], itd.

Prvi korak u realizaciji pseudoslučajnog pozicionog enkodera jeste nanošenje pseudoslučajne binarne sekvence (PRBS) na kodnu traku pri čemu je sekvenca prethodno generisana uz pomoć pomeračkog registra sastavljenog od n flip flopova i odgovarajuće povratne sprege definisane na osnovu tabele primitivnih polinoma. Paralelno, na sve stepene registra (flip-flopove) se dovodi signal takta i naredno stanje pomeračkog regista zavisi od prethodnog stanja i od definisane povratne sprege. Međutim, prilikom generisanja pseudoslučajne sekvence nije dozvoljeno pojavljivanje stanja u kome su svi stepeni pomeračkog registra na nuli, pošto generator ne može da izađe iz tog stanja. Primenjuje se linearna povratna sprega koja definiše ulazni bit pomeračkog registra kao linearnu funkciju prethodnog stanja registra. Naziv "linearna" potiče od činjenice da je povratna sprega izvedena primenom linearnih operacija, npr. sabiranjem po modulu 2, nad sadržajem memorijskih elemenata pomeračkog registra. Samo povratna sprega definisana na osnovu tabele primitivnih polinoma, [7], generisaće na izlazu pomeračkog registra sekvencu maksimalne dužine 2^n-1. Ovako generisana pseudoslučajna binarna sekvenca rezolucije n je maksimalne dužine 2^n-1 i sadrži 2^n-1 različitih kodnih reči dužine n. Preciznost merenja pozicije u pokretnim sistemima je veća ukoliko je veća rezolucija n. Kada je reč o industrijskim pokretnim sistemima sve više se povećavaju zahtevi što se tiče tačnosti i pouzdanosti pozicionog enkodera, a rezolucija koja se zahteva je reda $n \rangle 10$. U radu, na slici 1. prikazano je generisanje pseudoslučajne binarne sekvence maksimalne dužine, za $n = 10$ i odgovarajuće povratne sprege dobijene primenom primitivnog polinoma $h(x) = x^{10} + x^3 + 1$.



Sl. 1. Generisanje pseudoslučajne binarne sekvence za n=10.

Obrtni optički enkoderi visoke rezolucije, kao i linearni pozicioni enkoderi mogu da koriste metod pseudoslučajnog kodiranja. Određivanje pozicije kod ovih enkodera se zasniva na "osobini prozora" pseudoslučajne binarne sekvence [7], pri čemu prozor dužine n koji se kreće duž pseudoslučajne binarne sekvence izdvaja jedinstvenu kodnu reč. Susedne kodne reči koje su raspoređene na kodnoj traci se međusobno razlikuju samo u jednom bitu, što je omogućilo serijsko očitavanje koda. Međutim, serijsko očitavanje pseudoslučajnog koda zahteva početno inicijalno kretanje prilikom startovanja enkodera kako bi se formirala prva

Ivana Ranđelović – Elektronski fakultet, Univerzitet u Nišu, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: ivana.randjelovic@elfak.ni.ac.rs).
Dragan Denić – Elektronski fakultet, Univerzitet u Nišu, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: dragan.denic@elfak.ni.ac.rs).
Goran Miljković – Elektronski fakultet, Univerzitet u Nišu, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: goran.miljkovic@elfak.ni.ac.rs).

validna kodna reč, zbog čega su takvi enkoderi poznati na tržištu pod imenom “virtuelni enkoderi”. Kako bi dizajnirali enkoder koji je potpuno apsolutni u svim segmentima svog funkcionisanja, u ovom radu se uvodi paralelno očitavanje pseudoslučajnog koda. Odnosno, cilj je da se informacija o apsolutnoj poziciji dobija odmah i prilikom samog startovanja enkodera, i prilikom nestanka napajanja enkodera kao i u toku obrtanja osovine enkodera.

U prvom delu rada je objašnjen princip paralelnog očitavanja pseudoslučajnog koda korišćenjem linearnog niza fotodetektora. Pri tome je objašnjena procedura određivanja apsolutne pozicije, kao i problemi koji se pri tome javljaju. Istaknuta su ograničenja i prednosti korišćenja ovakvog načina funkcionisanja pseudoslučajnog enkodera. U drugom delu rada je predstavljena simulacija očitavanja koda korišćenjem LabVIEW okruženja, kao i simulacija rada elektronskog bloka ovakvog enkodera po prijemu signala iz sistema za očitavanje koda. Eksperimentisano je sa različitim putanjama kretanja, kao i sa različitim situacijama koje se mogu javiti pri funkcionisanju enkodera u industrijskom okruženju. Takođe, prikazana je i mogućnost primene linearnog niza fotodetektora, koji se trenutno mogu naći na tržištu, za paralelno očitavanje pseudoslučajnog koda na realizovanom staklenom kodnom disku.

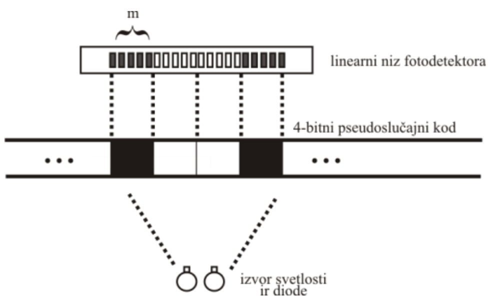
II. Paralelno očitavanje pseudoslučajnog koda kod pozicionih enkodera na bazi linearnog niza fotodetektora

Primena pozicionog pseudoslučajnog enkodera u aplikacijama gde je neprihvatljivo postojanje početnog inicijalnog kretanja kako bi se formirala prva apsolutna pozicija je dovelo do ideje primene metode paralelnog očitavanja pseudoslučajnog koda. Paralelno očitavanje pseudoslučajnog koda se može realizovati pojedinačnim fotodiodama, ali je njihova fotoosetljiva oblast reda 0.3 mm što sa kućištem bude 4-5 mm, pri čemu se značajno ograničava realizacija enkodera visoke rezolucije. Veće rezolucije se mogu postići primenom detektora sa integrisanim nizom fotodetektora na jednom čipu [8, 9], ili primenom CCD senzora [10]. Brzina očitavanja CCD senzora je manja od brzine očitavanja niza fotodetektora, pa se CCD senzori mogu koristiti za enkodere u aplikacijama gde imamo manje brzine rotacije osovine.

Na tržištu se mogu naći komercijalni integrisani linearni nizovi fotodetektora sa različitim širinama fotoosetljivog elementa, odnosno različitim razmacima između susednih fotodetektora koji su reda 400 µm, 100 µm, 50 µm, 25 µm itd. Neki od prozvođača ovakvih linearnih nizova fotodetektora su: iC-Haus GmbH [14], Hamamatsu company, [15], Sensors Unlimited [16].

Osnovni princip funkcionisanja pseudoslučajnog enkodera sa paralelnim očitavanjem na bazi primene linearnog niza fotodetektora je prikazan na slici 2. i sastoji se u sledećem.

Primenjuje se kodni disk transparetnog tipa gde se sa jedne strane pseudoslučajne kodne trake nalazi izvor svetlosti, infracrvene diode, a sa druge strane kodne trake integrisano kolo sa linearnim nizom fotodetektora. Za određivanje apsolutne pozicije potrebno je očitati *n* uzastopnih bitova sa pseudoslučajne kodne trake.



Sl. 2. Primer korišćenja linearnog niza fotodetektora za očitavanje 4-bitnog pseudoslučajnog koda

Kodna traka je tako projektovana da je širina bita, zbog pouzdanosti očitavanja, značajno veća od širine fotoelementa linearnog niza fotodetektora. Odnosno, jedan bit sa pseudoslučajne kodne trake čitaju m susednih fotodetektora. Na primer, ako je $m=5$ u idealnom slučaju dobili bismo ovakvo očitavanje kodne reči $\{1001\}$ sa 4-bitne pseudoslučajne kodne trake pomoću linearnog niza fotodetektora $\{\ldots 11111000000000011111\ldots\}$. Može se uočiti da je broj uzastopnih jedinica i uzastopnih nula po jednom bitu upravo *m*. U realnom sistemu, broj uzastopnih jedinica i nula će varirati od broja *m*, $\{\ldots 111111000000000011110\ldots\}$, zbog vibracija, temperature, prašine, itd. Kako je poznat razmak između susednih bitova, određuje se gruba apsolutna pozicija prikupljanjem susednih *n* bitova i njihovom konverzijom iz pseudoslučajnog u prirodni kod.

Očitavanja pseudoslučajnog bita koja se prikupljaju sa krajnje desnim fotodetektorima omogućavaju dodatnih (m-1) bitova koji se mogu iskoristiti za određivanje fine pozicije pokretnog sistema. Time bi se mogla povećati rezolucija pseudoslučajnog pozicionog enkodera. Takođe, smer kretanja pokretnog sistema se može odrediti praćenjem promena vrednosti fine pozicije. Gruba pozicija se određuje uvek kada se javi skok fine pozicije. Što se tiče određivanja smera kretanja pokretnog sistema može se zaključiti da prilikom kretanja pokretnog sistema udesno, fina pozicija skače od vrednosti *"fina pozicija=m-1"* na *"fina pozicija=0"*. Kada je skok detektovan, fina pozicija dobija vrednost 0, i sledi određivanje grube pozicije. U slučaju kretanja pokretnog sistema ulevo, fina pozicija skače od vrednosti *"fina pozicija=0"* na *"fina pozicija=m-1"*. Gruba pozicija se određuje kada je vrednost *"fina pozicija=m-1"* i smanjuje se za jedan.

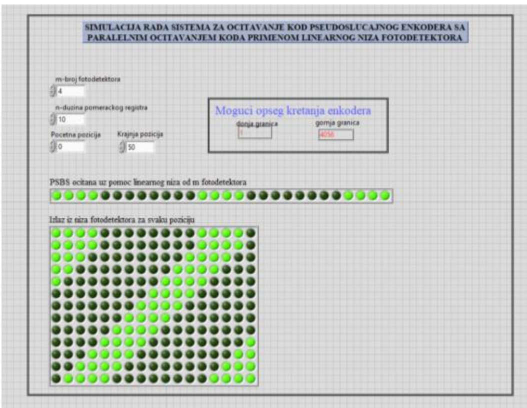
Algoritmi rada enkodera na osnovu kojih je i urađena digitalna simulacija rada predloženog enkodera detaljno su opisani u referencama [11, 12] i omogućavaju opsežna istraživanja razlilitih varijanti rešenja enkodera pre njihove praktične realizacije. Algoritam u referenci [11] je u kasnijim istraživanjima modifikovan u pogledu dobijanja jednostavnijeg rešenja i kao takav prikazan u referenci [12] pri čemu je pokazano da u predloženom rešenju greške koje uzrokuju skok fine pozicije ne utiču na performanse sistema.

III. Digitalna simulacija rada predloženog rešenja enkodera sa paralelnim očitavanjem pseudoslučajnog koda

U cilju poboljšanja funkcionalnosti i brzine algoritma potrebno je izvršiti veći broj eksperimenata. Istraživanja su vršena na polju gde realna proba određenih rešenja zahteva velika ekonomska sredstva. Razvoj pozicionog pseudoslučajnog enkodera korišćenjem realnih komponenti enkodera bi bio skup i vremenski duži postupak, zbog čega se primenjuje računar i LabVIEW okruženje u cilju dobijanja raznih varijanti rešenja enkodera na brži i jeftiniji način. U tom pogledu, za istraživanje je najbolje posedovati dobar eksperimentalni sistem. Kvalitetnim programima za simulaciju kretanja pokretnog sistema, generisane su što tačnije informacije koje bismo dobili na izlazima sistema za očitavanje kod realnog enkodera. Dobijene informacije se zatim koriste kao ulaz programa koji simulira rad elektronskog bloka i algoritma rada enkodera. Programi za simulaciju rada enkodera su realizovani primenom softverskog paketa LabVIEW, [13].

Simulacija počinje startovanjem prvog programa koji generiše niz digitalnih signala koji se upisuju u datoteku, odakle se mogu očitati drugim programom za simulaciju rada elektronskog bloka enkodera. U samoj izradi simulatora enkodera najpre se krenulo od pretpostavke da sistem radi idealno, odnosno bez grešaka u očitavanju koda. Zatim su u simulator implementirani različiti ometajući faktori koji bi se javili u realnim industrijskim uslovima i određeno je kako to utiče na izlazne signale simulatora.

Dakle, u okviru razvijenog rešenja, dat je primer za pseudoslučajnu sekvencu sa potrebnim brojem bitova $n = 10$, broj fotodetektora po jednom bitu koda je $m = 4$, i zadati opseg kretanja $(0 - 50)$, slika 3.

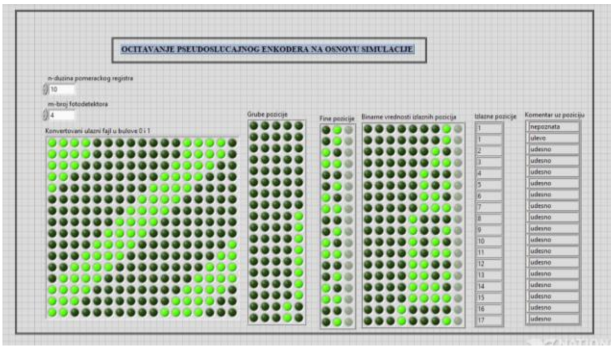


Sl. 3. Front panel simulatora rada sistema za očitavanje kod pseudoslučajnog pozicionog enkodera sa paralelnim očitavanjem koda

Podaci se upisuju u datoteku a zatim se startovanjem drugog programa određuje pozicija pokretnog sistema, slika 4.

Trebalo bi naglasiti da su grube i fine pozicije date u dekadnom sistemu, čije se vrednosti prevode u binarni oblik. Na vrednost grube pozicije u binarnom obliku dopisuje se binarna vrednost fine pozicije i prevođenjem u dekadni sistem dobija se trenutna vrednost izlazne pozicije. Na primer: ako je “gruba pozicija=7”, “fina pozicija=0”, binarne vrednosti su “gruba pozicija=111 i “fina pozicija=000”. Izlazna pozicija je sada binarno “111000” što bi u dekadnom sistemu odgovaralo poziciji $p = 56$. Smer kretanja pokretnog sistema jeste u formi “komentar uz poziciju”.

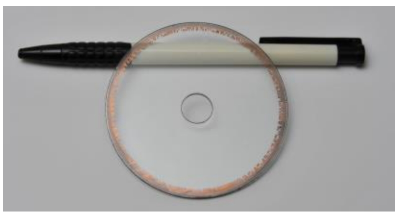
Velike su prednosti ovakvog realizovanog sistema. Testiranje određenog rešenja je veoma slično kao i kod realnog pozicionog enkodera. Rezultati simulacije se mogu prikazati ili odštampati što obezbeđuje brze korekcije eventualnih grešaka u testiranom rešenju.



Sl. 4. Front panel simulatora elektronskog bloka pseudoslučajnog pozicionog enkodera sa paralelnim očitavanjem koda

IV. Analiza metode paralelnog očitavanja pseudoslučajnog koda primenom linearnog niza fotodetektora kod predloženog enkodera

U cilju dalje praktične provere predloženog rešenja enkodera za koje je urađena digitalna simulacija korišćen je stakleni disk enkodera i data je analiza primene linernog niza fotodetektora konkretnih proizvođača. Kodni disk je izrađen od optičkog stakla, odgovarajućeg kvaliteta obrade, debljine 1,6 mm i prečnika 72,7 mm. Prečnik otvora u centru kodnog diska iznosu 12 mm. Na površinu kodnog diska, koja je polirana do traženog kvaliteta obrade, nanosi se kodna traka u vidu preciznih podela, odnosno šara, sačinjenih od prozračnih i neprozračnih polja, kao što je prikazano na slici 5. Rezolucija pseudoslučajne trake je 10 bita, odnosno svaka pseudoslučajna traka sadrži 1024 polja. Polje jednog bita na pseudoslučajnoj kodnoj traci je širine 201 µm i visine 800 µm.



Sl. 5. Fotografija realizovanog staklenog diska pseudoslučajnog pozicionog enkodera

Kodni disk je urađen za serijsko očitavanje pseudoslučajnog koda, poseduje dodatnu pseudoslučajnu kodnu traku i sinhronizacionu traku, koje nisu potrebne za analizu ovde predloženog rešenja očitavanja koda.

Za analizu očitavanja pseudoslučajnog binarnog koda je izabran senzor, LC/LSC serije, proizvođača Sensors Unlimited [16], čitač sa linearnim nizom veoma brzih InGaAs fotodetektora pri čemu je razmak između fotodetektora 25 μm dok je broj fotodetektora 512. Analizom širine kodnog bita i širine fotodetektora ustanovljeno je da je za očitavanje jednog bita sa pseudoslučajne kodne trake potrebno $m = 4$ fotodetektora. Odnosno, za očitavanje 10-bitnog pseudoslučajnog koda bi se koristilo 40 susednih fotodetekora. Na osnovu kataloških podataka, brzina ovih senzora je 91,912 kHz dok je vreme očitavanja 10,88 μs pri takt frekvenciji od 12,5 MHz. Ukoliko bi se koristio integrisani niz fotodetektora drugog proizvođača, konkretno InGaAs G7151-16 proizvođača Hamamatsu company [15], koji sadrži 16 fotodetektora i razmak između senzora iznosi 100 μm tada bi bilo potrebno $m = 2$ fotodetektora da bi se očitao jedan bit sa pseudoslučajne kodne trake na staklenom disku enkodera.

Istraživanja vezana za paralelno očitavanje pseudoslučajnog koda na bazi linearnog niza fotodetektora kod pozicionog enkodera su samo započeta. Primenom konkretnih integrisanih nizova fotodetektora poznatih proizvođača upotrebljenih za paralelno očitavanje pseudoslučajnog koda otvara se mogućnost i praktične realizacije enkodera kao pozicionog enkodera opšte namene visoke rezolucije.

V. Zaključak

Danas, sve je više prisutan zahtev u industriji za pouzdanim i tačnim informacijama o poziciji pokretnog sistema. Za potrebe razvoja pseudoslučajnih pozicionih enkodera implementirana su dva programa primenom softverskog paketa LabVIEW pri čemu jedan program simulira sistem za očitavanje koda pseudoslučajnog pozicionog enkodera, dok drugi program simulira funkcionisanje elektronskog bloka takvog enkodera. Predstavljena su neka trenutno dostupna kola lineranih nizova fotodetektora na tržištu i data je analiza primene nekih od njih za predloženo rešenje enkodera. Potrebno je još dosta eksperimentisanja kako bi se došlo do konačnog praktičnog rešenja ovakvog enkodera, pri čemu bi on mogao da ima niz prednosti u odnosu na neka prethodna rešenja u literaturi. Predloženi enkoder bi bio apsolutni enkoder gde se ne zahteva početno inicijalno kretanje i koji omogućuje primenu metoda za povećanje njegove rezolucije korišćenjem informacija o finoj poziciji sistema. Sa druge strane, kao nedostatak ovako realizovanog rešenja bi bila veličina senzora kao i sporije očitavanje senzora.

Zahvalnica

Ovaj rad je podržan od strane Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije.

Literatura

[1] D. Denić, G. Miljković, “Code reading synchronization method for pseudorandom position encoders”, Sensor. Actuat. A-Phys. vol. 150, pp. 188-191, 2009.

[2] D. Denić, I. Stojković, “Pseudorandom/natural code converter with parallel feedback logic configuration”, Electron. Lett. vol. 46, pp. 921-922, 2010.

[3] D. Denić, M. Arsić, “Checking of pseudorandom code reading correctness”, Electron. Lett. vol. 29, pp. 1843-1844, 1993.

[4] D. Denić, I. Ranđelović, G. Miljković, “Recent trends of linear and angular pseudorandom encoder development”, Proceedings of International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM), Taormina, Sicily, Italy, pp. 746-750, 2006.

[5] I. Stojković, G. Miljković, D. Denić, D. Živanović, “Application of pseudorandom position encoder for crane positioning”, Proceedings of the International Scientific Conference UNITECH 2017, Gabrovo, Bugarska, vol. 3, pp. 289-292, 2017.

[6] E.M. Petriu, J.S. Basran, F.C.A. Groen, “Automated guided vehicle position recovery”, in IEEE Trans. Instrum. Meas. vol. 39, pp. 254-258, 1990.

[7] F.J. MacWilliams, N.J.A Slone, “Pseudo-random sequences and arrays”, Proceeding of IEEE, Vol. 64, No. 12, pp. 1715-1728, 1976.

[8] J.T.M Stevenson and J.R. Jordan, “Absolute position measurement using optical detection of code patterns”, J. Phys E. Sci. Instrum. 21, pp. 1140-1145, 1988.

[9] M. Durana, P. Gallay, P. Robert, F.C. Pruvot, “Novel type submicrometre resolution pseudorandom position optical encoder”, Electronic. Letters., Vol 29, No 20, pp. 1792-1794, 1993.

[10] H. Khalfallah, E.M. Petriu, F.C.A. Groen, “Visual position recovery for an automated guided vehicle”, IEEE Trans. Instrum. and Meas., Vol. 41, No. 6, pp. 906-910, 1992.

[11] D. Denić, M. Rančić, G. Miljković, “Pseudoslučajni apsolutni pozicioni encoder visoke rezolucije”, Zbornik radova 47. Konferncije ETRAN, str. 452-455, 2003.

[12] D. Denić, I. Ranđelović, M. Rančić, “High-resolution pseudorandom encoder with parallel code reading”, Electronics and electrical engineering Nr 7(56), Kaunas, Litvania, pp. 9-14, 2004.

[13] LabVIEW, *User's Manual*, National Instruments, www.ni.com, USA, 2010.

[14] www.ichaus.com

[15] www.hamamatsu.com

[16] www.sensorsinc.com

Abstract

The paper proposes the application of linear array of photodetectors for parallel code reading of pseudorandom binary code in absolute pseudorandom position encoders. The proposed method of code reading would eliminate the need for the initial movement of the shaft of encoder during starting in the case of serial code reading. A simulation of the proposed solution of position encoder using the LabVIEW software package was performed. Digital simulation was realized in the form of two programs, where one program simulates a code reading system of a pseudorandom position encoder, while the other program simulates the operation of an electronic block of such an encoder. An analysis of the application of a commerrcially available linear array with photodetectors for code reading, realized from glass disk of an encoder with pseudorandom code track, is presented.

A code reading based on linear array of photodetectors at pseudorandom position encoders

Ivana Ranđelović, Dragan Denić i Goran Miljković

Snimanje UI karakteristike odvodnika prenapona

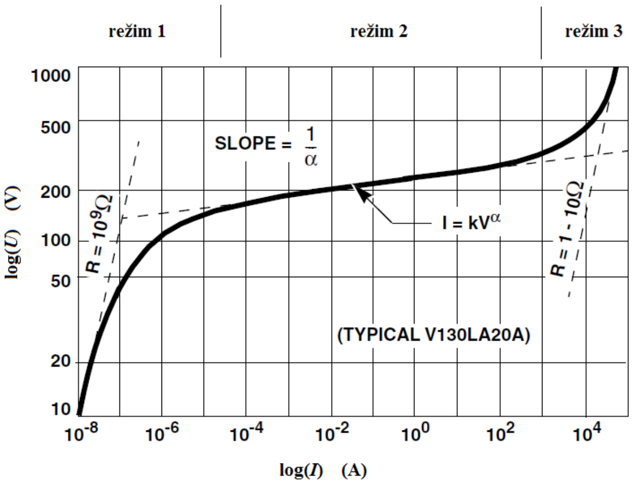
Dragan Pejić, Boris Antić, Zoran Mitrović, Nemanja Gazivoda, Marina Subotin

***Apstrakt*—Ovaj rad daje prikaz iskustava prilikom snimanja naponsko-strujne (UI) karakteristike odvodnika prenapona (ZnO varistor). Snimanje je obavljeno korišćenjem dvokanalnog digitalnog osciloskopa. Napon sa krajeva varistora je dovođen na jedan kanala osciloskopa, a drugi kanal je iskorišćen za određivanje struje kroz varistor merenjem napona na redno vezanom šantu. Analiza je obavljena na dva načina: a) posmatranjem efektivne vrednosti napona i struje i b) posmatranjem trenutne vrednosti napona i struje varistora. Na kraju su dobijeni razultati i primenjene metode upoređene sa oskudnim informacijama dostupnim u literaturi.**

***Ključne reči*—naponsko-strujna karakteristika, efektivna vrednost, odvodnik prenapona, varistor.**

I. Uvod

Varistor je nelinearni dvokrajni element sa naponsko-strujnom (UI) karakteristikom pogodnom za zaštitu elektronskih kola, energetskih postrojenja i dalekovoda od prenaponskih pojava. UI karakteristika varistora [1] se može podeliti u tri zone, Sl. 1.



Sl. 1. Naponsko-strujna (UI) karakterstika varistora, sa obeleženim režimima rada

Kada je napon manji od napona prvog "kolena" reč je o režimu velike otpornosti, tako da kroz varistor protiče mala struja reda 10^{-8} A do 10^{-3} A (režim curenja). S obzirom da je napon manji od kritične vrednosti, varistor se "trudi" da bude neprimetan; idealno bi bilo da je struja kroz varistor jednaka nuli. U ovom režimu na ponašanje varistora značajno utiče ekvivalentna kapacitivnost, zbog koje, u prosto-periodičnom kolu, dolazi da faznog pomeranja između napona i struje za približno 90˚.

U drugom režimu rada varistor ima mnogo manju dinamičku otpornost (idealno bi bilo da je dinamička otpornost jednaka nuli), tako da porastu napona odgovara višestruko povećanje struje. U ovom režimu rada, varistor ograničava prenapon konzumiranjem vrlo velikih struja, reda kA (normalan režim rada varistora). Ako je reč o elektrostatičkim pražnjenjima, usled konzumiranja vrlo velike struje se očekuje da električna energija u kratkom vremenskom intervalu bude potrošena, te da se varistor vrati u prvi režim rada.

Na kraju normalnog režima rada se nalazi drugo "koleno" karakteristike varistora. Za napone veće od napona drugog kolena, nastaje treći deo UI karakteristike (režim preokreta). U trećem režimu rada varistor povećava svoju otpornost, tako da sa daljim porastom napona struja sporije raste. Ukoliko ne bi bilo ovog režima, postojala bi opasnost da se, usled razvijanja velike disipacije, ošteti varistor ili štićeni objekat.

Na Sl. 1 je horizontalna osa prikazana u logaritamskoj raspodeli, kako bi se dobio pregledan grafik za veliki opseg struje od 10^{-8} A do 10^{5} A.

II. Laboratorijski eksperiment

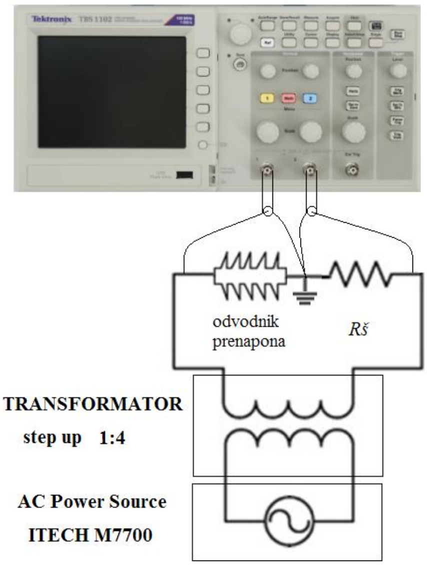
Na Sl. 2 je prikazana šema koja je korišćena za snimanje UI karakteristike varistora B72232D 271K161 [2]. Ovaj varistor je namenjen za zaštitu od prenapona niskonaponske mreže 230 V. Napon od 275 V je napon "kolena" koji razgraničava režim curenja i normalan režim rada varistora.

Naponski izvor maksimalne snage 300 W i maksimalnog napona 300 V proizvođača ITECH [3] napaja transformator prenosnog odnosa 1:4. Transformator napaja rednu vezu ispitivanog varistora i rednog otpornika - šanta $R_š$. Dvokanalni digitalni osciloskop [4] je sondama odgovarajućeg faktora slabljenja povezan tako da se na jedan kanal dovodi napon sa varistora, a na drugi kanal napon sa šanta. Poznavanjem otpornosti šanta i primenom Omovog zakona određuje se struja kroz rednu vezu šanta i varistora.

Osciloskop je korišćen na dva načina. U prvoj situaciji je podešeno da osciloskop određuje efektivnu vrednost jednog i drugog napona. Za razne vrednosti prostoperiodičnog napona je beležena efektivna vrednost napona varistora, a primenom Omovog zakona je određivana efektivna vrednost struje na osnovu efektivne vrednosti napona i poznavanja otpornosti šanta. Budući da se snimanje UI karakteristike vrši u naizmeničnom režimu, autorima se učinilo logičnim

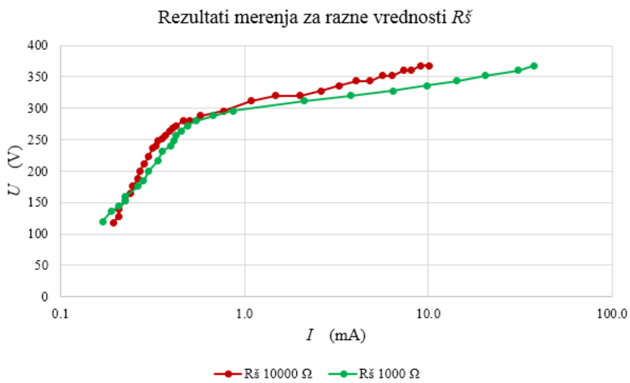
Dragan Pejić, Univerzitet u Novom Sadu, Fakultet tehničkih nauka, Dositeja Obradovića 6, Novi Sad (e-mail: pejicdra@uns.ac.rs).
Boris Antić, Univerzitet u Novom Sadu, Fakultet tehničkih nauka, Dositeja Obradovića 6, Novi Sad (e-mail: antic@uns.ac.rs).
Zoran Mitrović, Univerzitet u Novom Sadu, Fakultet tehničkih nauka, Dositeja Obradovića 6, Novi Sad (e-mail: zoranmit@uns.ac.rs).
Nemanja Gazivoda, Univerzitet u Novom Sadu, Fakultet tehničkih nauka, Dositeja Obradovića 6, Novi Sad (e-mail: nemanjagazivoda@uns.ac.rs).
Marina Subotin, Univerzitet u Novom Sadu, Fakultet tehničkih nauka, Dositeja Obradovića 6, Novi Sad (e-mail: marina.bulat@uns.ac.rs).

da se vrši beleženje efektivne vrednosti napona na varistoru, kao i efektivna vrednost napona na šantu na osnovu koje je računata efektivna vrednost struje kroz varistor. „Step-up“ transformator je korišćen zbog nedovoljnih naponskih mogućnosti izvora napona ITECH M7700.



Sl. 2. Šema korišćena za snimanje UI karakteristike varistora primenom dvokanalanog osciloskopa

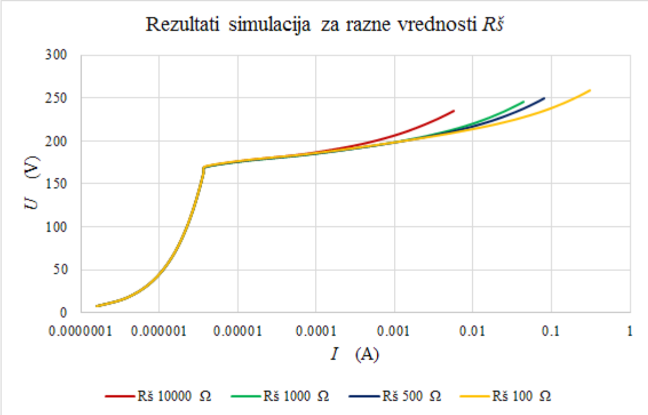
Eksperiment je ponavljan za dve vrednosti šanta: 1 kΩ i 10 kΩ. Ideja je bila da se vrednosti šanta prilagode redu veličine struje koja se želi meriti, kako bi se dobijao napon na šantu merljiv primenom osciloskopa. Na Sl. 3 su prikazani rezultati merenja dobijeni za pomenute vrednosti šanta.



Sl. 3. Rezultati merenja na osnovu efektivne vrednosti napona i struje varistora

Dobijeni rezultati, na prvi pogled deluju zbunjujuće, jer se čini da se UI karakteristika varistora menja u zavisnosti od vrednosti redno vezanog šanta. Napominjemo da je na merenje primenjena procedura koja je interpretirana na osnovu preporuka međunarodnih standarda za ispitivanje UI karakteristike odvodnika, te da su sve greške proizišle iz dvosmislenosti ili nedorečenosti koje postoje u ovim standardima.

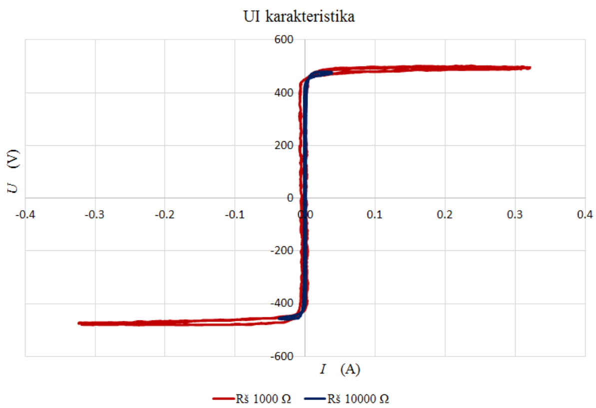
Da bi se eliminisao uticaj grešaka merenja, urađene su simulacije u paketu LT Spice [5]. Na Sl. 4 su prikazani rezultati dobijeni simulacijama, uz verno oponašanje opisanog načina merenja. U nedostatku modela ispitivanog varistora, korišćen je model varistora manjeg napona kolena, što ne umanjuje ispravnost zaključaka koji se iz rezultata simulacija mogu izvesti.



Sl. 4. Rezultati simulacija u paketu LT Spice na osnovu efektivne vrednosti napona i struje varistora

Na osnovu rezultata simulacija može se zaključiti da i simulacije potvrđuju da ovako dobijena UI karakteristika varistora zavisi od vrednosti redno vezanog šanta. Rezultati simulacija daju UI karakteristiku pravilnijeg oblika u poređenju sa UI karakteristikom dobijenom na osnovu merenja. Objašnjenje leži u uticaju grešaka koje postoje u procesu merenja, a ne postoje u procesu simuliranja. Greške se mogu pripisati ugrađenim AD konvertorima u digitalni osciloskop, kao i načinu računanja efektivne vrednosti na osnovu odbiraka napona.

Snimanje UI karakteristike je obavljeno i na drugi način. Kratko, u trajanju od nekoliko sekundi, je na rednu vezu varistora i šanta doveden prostoperiodični napon maksimalne efektivne vrednosti. Pod ovim se misli na maksimalnu vrednost koju je ukupna aparatura prikazana na Sl. 2 bila u stanju da da, kako po pitanju napona, tako i po pitanju struje. Osciloskop je konfigurisan da radi u režimu „one shot“. Na ovaj način su autori pokušavali da postignu najveće moguće napone i struje varistora, a da izbegnu pregrevanje i eventualno oštećivanje varistora. Ovog puta je pristup merenju bio drugačiji. Osciloskop nije korišćen kao merilo efektivne vrednosti dva napona. U memoriju osciloskopa su snimljeni odbirci napona varistora i šanta. Na osnovu odbiraka napona šanta su, primenom Omovog zakona, određene trenutne vrednosti struje kroz rednu vezu varistora i šanta.



Sl. 5. UI karakteristika varistora dobijena na osnovu odbiraka napona i struje

Na Sl. 5 je prikazana UI karakteristika varistora dobijena ovim postupkom. Budući da vrednosti odbiraka mogu biti i pozitivne i negativne (za razliku od efektivne vrednosti koja ne može biti negativna), grafik na Sl. 5 je prikazan u sva četiri kvadranta. Zbog činjenice da su u eksperimentu postignute male vrednosti struje, dovoljne tek da se pređe prvo koleno karakteristike, nije bilo potrebe za logaritamskom horizontalnom osom, već je korišćena linearna osa za struju. Primetno je da UI karakteristika prikazana na Sl. 5 ne zavisi od vrednosti korišćenog šanta.

U svim prethodnim situacijama, isključivo zbog ograničenja i mogućnosti raspoložive opreme, izvršeno je snimanje u prva dva režima: u režimu curenja i normalnom režimu rada. Nije vršeno snimanje u režimu preokreta, zbog nemogućnosti raspoložive opreme da obezbedi dovoljno veliku struju za ispitivanje varistora, a i zbog činjenice da se u ovom režimu rada ispitivani varistor nađe samo u ekstremnim situacijama, nakon kojih najčešće biva nepovratno oštećen i isključen iz zaštitnog kola.

III. Diskusija

Omov zakon u kolu vremenskih promenljivih napona i struja, a za otpornost R dat je (1).

$$u(t) = R \cdot i(t) \tag{1}$$

Sada možemo odrediti srednju kvadratnu vrednost leve i desne strane jednakosti.

$$\frac{1}{T}\int_0^T u^2(t)\,dt = \frac{1}{T}\int_0^T R^2 \cdot i^2(t)\,dt \tag{2}$$

Kada je otpornost konstantna i nezavisna od vremena i struje koja kroz nju protiče (a time i od napona na njenim krajevima), a imajući u vidu definiciju efektivne vrednosti, kvadrat otpornosti možemo izvući ispred integrala, onda dobijamo (3). Jednačina (3) pokazuje da se vrednost otpornosti može dobiti deljenjem efektivne vrednosti napona na krajevima otpornika i efektivne vrednosti struje koja protiče kroz otpornik.

$$U_{ef}^2 = R^2 \cdot I_{ef}^2 \quad \Rightarrow \quad R = U_{ef} / I_{ef} \tag{3}$$

U slučaju varistora, kao nelinearnog otpornog elementa, situacija je komplikovanija. Ovoga puta otpornost nije konstantna, nego zavisi od napona i struje. U tom slučaju (2) poprima drugačiji oblik.

$$\frac{1}{T}\int_0^T u^2(t)\,dt = \frac{1}{T}\int_0^T R^2(t) \cdot i^2(t)\,dt \tag{4}$$

Važno je primetiti da sada nije moguće iz integrala sa desne strane (4) izvući kvadrat otpornosti R, jer zavisi od struje i, a struja zavisi od vremena t. Iz ovoga sledi da nije moguće dobiti izraz koji odgovara (3) u slučaju linearnog otpornika. To znači da određivanje otpornosti na osnovu količnika efektivne vrednosti napona i efektivne vrednosti struje daje ispravan rezultat samo ako je posmatrana otpornost linearna. S druge strane, ovaj postupak bi bio pogrešan ako bismo ga primenjivali u slučaju nelinearne otpornosti. To je ujedno i osnovni uzrok greške u interpretaciji rezultata, usled načina na koji se UI karakteristike i preporučene šeme merenja prikazuju u standardima ili katalozima proizvođača varistora, jer je svuda indikovano merenje prostoperiodičnom pobudom uz napomenu da su sve vrednosti efektivne.

U prvom načinu snimanja UI karakteritike varistora (rezultati merenja prikazani na Sl. 3, a rezultati simulacija na Sl. 4) otpornost je određivana na osnovu efektivne vrednosti napona i efektivne vrednosti struje. U slučaju određivanja efektivne vrednosti struje na osnovu efektivne vrednosti napona na šantu, struja je određena korektno, jer je otpornost šanta konstantna i linearna. Propust je učinjen kod primene istog načina razmišljanja na varistor, jer je varistor vrlo nelinearan dvokrajni element. UI karakteristika snimljena na ovaj način zavisi od vrednosti šanta u normalnom režimu rada, gde je i nelinearnost izraženija, dok se u režimu curenja (u uslovima približne linearnosti) uticaj vrednosti šanta i ne primećuje, pogotovo u rezultatima dobijenim simulacijama. Ovo je objašnjenje naizgled paradoksalne pojave prikazane na Sl. 3 i 4, po kojima ispada da UI karakteristika zavisi od vrednosti korišćenog šanta.

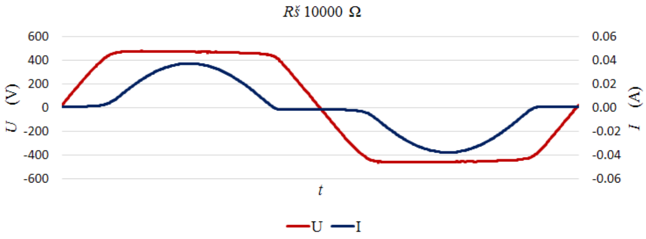
U drugom načinu snimanja UI karakteristike ne dolazi do ovakvog propusta jer je rađeno sa trenutnim vrednostima napona i struje.

Zbog čega je moguće da u interpretaciji preporuke za ispitivanje UI karatkeristike varistora dođe do ovakve zabune?

U slučaju primene varistora kao odvodnika prenapona u niskonaponskoj mreži, napon na krajevima varistora je naizmenični, približno prostoperiodičnog talasnog oblika. Proizvođači daju naponsko-strujnu karakteristiku varistora u prvom kvadrantu, dakle za pozitivne vrednosti napona i struja. Ovo može biti obrazloženo na nekoliko načina: a) zbog simetrije karakteristike dovoljno je prikazivati ponašanje u prvom kvadrantu (za pozitivne vrednosti napona i struje), odnosno nema potrebe da se prikazuje i situacija kada napon i struja imaju negativne vrednosti, b) naponsko-strujna karakteristika je snimljena na osnovu merenja efektivne vrednosti napona i efektivne vrednosti struje, a poznato je da su efektivne vrednosti nenegativne.

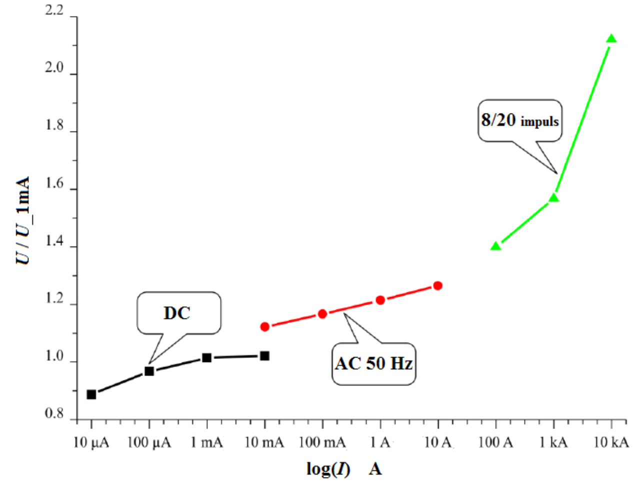
Na osama UI karakteristike su date efektivne vrednosti napona i struje, jer su odvodnici napona namenjeni za korišćenje u naizmeničnom režimu, gde se uglavnom operiše efektivnim vrednostima napona i struja mnogo češće nego trenutnim vrednostima.

Na Sl. 5 je prikazana UI karakteristika varistora snimljena na osnovu trenutnih vrednosti napona i struja. Zato se UI karakteristika proteže u prvom i trećem kvadrantu UI ravni. Zbog toga što su na osama prikazane trenutne vrednosti napona i struje, dobijeni grafik se po obliku slaže, ali se po vrednostima ne slaže sa specifikacijom proizvođača. Proizvođač je napone i struje na osama naveo u formi efektivnih vrednosti. Tek kada se vrednosti napona i struje na osama na Sl. 5 deljenjem faktorom $\sqrt{2}$ ponderišu, dobija se slaganje snimljene UI karakteristike varistora sa specifikacijom proizvođača. Pomenuti faktor $\sqrt{2}$ predstavlja odnos amplitude i efektivne vrednosti za prostoperiodičan talasni oblik, a sa Sl. 6 se jasno vidi da ni struja ni napon nemaju podrazumevani prostoperiodični oblik.



Sl. 6. Oblik napona i struje varistora pri korišćenju šanta od 10000 Ω

Daljim proučavanjem slabo dostupne i nejasne dokumentacije, autori su došli do saznanja da su i drugi imali slične probleme prilikom određivanja UI karakteristike varistora. Kao primer ovde ćemo navesti UI karakteristiku sa Sl. 7 preuzetoj iz [6], na kojoj su naznačene tri zone. U svakoj od zona je otpornost određena drugačijim postupkom. Razlog za ovo jeste izuzetno veliki opseg struja za koje je potrebno ispitati varistor, što za posledicu ima veliku snagu, odnosno energiju koju bi izvor trebalo da proizvede tokom ispitivanja.



Sl. 7. UI karakteristika snimljena na tri načina, u DC, AC i impulsnom režimu

Na Sl. 7 se vidi da postoje diskontinuiteti u snimljenoj UI karakterstici. Za najmanje vrednosti struja, koje se javljaju u režimu curenja, predlaže se ispitivanje varistora u jednosmernom (DC) režimu. Varistor je tada moguće modelovati paralelnom vezom vrlo velike otpornosti (približno linearne u režimu curenja) i malom kapacitivnošću. Merenjem u DC režimu se uticaj kapacitivnosti svesno zanemaruje. U normalnom režimu rada varistora se predlaže snimanje UI karakteristike primenom napona mrežne učestanosti (AC režim). U trećem režimu, režimu preokreta, se predlaže korišćenje impulsa definisanog oblika [6], čija uzlazna ivica traje 8 µs, a silazna 20 µs.

IV. Zaključak

U radu su prikazana dva načina snimanja UI karakteristike ZnO varistora namenjenog za zaštitu niskonaponske mreže od prenapona. Na grafiku koji daju proizvođači su date efektivne vrednosti napona i struja. Problem nastaje kada se UI karakteristika varistora pokušava snimiti merenjem efektivne vrednosti napona i struje, jer se u tom slučaju dobija da UI karakteristika zavisi od vrednosti redno vezanog šanta korišćenog za merenje struje.

U radu je pokazano da se ispravan način snimanja UI karaketristike sprovodi beleženjem trenutnih vrednosti napona i struje. S obzirom da je u naizmeničnom režimu rada uobičajeno iskazivanje efektivnih vrednosti mnogo češće nego trenutnih vrednosti, konačan izgled UI karakteristike se dobija pravilnim ponderisanjem trenutnih vrednosti napona i struja radi dobijanja efektivnih vrednosti. Problem nastaje kada se pretpostavlja ponder vrednosti $\sqrt{2}$ (odnos amplitude i efektivne vrednosti u slučaju prostoperiodičnog talasnog oblika), a da pri tome ni struja ni napon nemaju podrazumevani prostoperiodični oblik.

Autori predlažu snimanje, ali i prikazivanje UI karakteristike varistora na osnovu trenutnih vrednosti napona i struja i bez dodatnog ponderisanja, čime će se dobijati verodostojniji rezultati, koji će biti ponovljivi, neće zavisiti od uslova merenja i moći će se nedvosmisleno interpretirati.

Zahvalnost

Ovo istraživanje je finansirano od strane projekta Fonda za inovacionu delatnost Republike Srbije, pod nazivom “New generation of silicone and zinc oxide surge arrester and insulators for low and medium voltage levels", broj projekta IF ID: 50205.

Literatura

[1] Morris, W., “Physical Properties of the Electrical Barriers in Varistors,” J. Vac. Sci. Technol., 13,1976, pp. 926-931.
[2] Metal Oxide Varistor, Round Disk Metallised, https://www.tdk-electronics.tdk.com/
[3] IT-M7700 High Performance Programmable AC Power Supply, https://www.itechate.com/en/product/ac-power-supply/IT-M7700.html
[4] Tektronix TBS1102 Digital Storage Oscilloscope, https://www.tequipment.net/TektronixTBS1102.html
[5] LT Spice, https://www.analog.com/en/design-center/design-tools-and-calculators/ltspice-simulator.html#
[6] Surge protective component application guide - metal oxide varistor (MOV) component

Abstract

This paper shows the experience of recording of a voltage-current (UI) characteristic of a surge arrester (ZnO varistor). A two channel digital oscilloscope was used for the recording. The voltage from the terminals of the varistor was fed into one of the channels of the oscilloscope, while the other channel was used to determine the value of current through the varistor by measuring voltage on a series shunt. The analysis included: a) observing the RMS values of current and voltage; and b) observing the instantaneous values of the varistor's voltage and current. This was followed by obtaining the results and comparing the methods applied with the scarce information that the literature offers.

The Recording of an UI Characteristic of a Surge Arrestor

Dragan Pejić, Boris Antić, Zoran Mitrović, Nemanja Gazivoda, Marina Subotin

Sortiranje predmeta prema boji akvizicijom videa primenom virtuelne instrumentacije

Branko Stojković, Branko Koprivica, Alenka Milovanović i Tatijana Dlabač

***Apstrakt*—Cilj ovog rada je prikaz laboratorijske postavke za sortiranje predmeta prema boji, akvizicijom videa u programskom paketu LabVIEW. U radu su date osnovne informacije o programskim dodacima koji su potrebni za rad sa videom i slikom u programu LabVIEW i njihove mogućnosti primene, a ukratko je opisan i HSL sistem konverzije slike. Princip rada postavke za sortiranje predmeta prema boji, programska kalibracija USB kamere i postupak izvođenja laboratorijske vežbe su takođe opisani u radu. Na kraju rada su prikazani rezultati merenja mase sortiranih predmeta i data je odgovarajuća diskusija.**

***Ključne reči*—Virtuelna instrumentacija; Akvizicija videa i slike; USB kamera; LabVIEW; Sortiranje.**

I. Uvod

Kako bi se odgovorilo trenutnim zahtevima koji su postavljeni u toku razvoja industrije i tehnologije potrebno je osmisliti nove ili unaprediti postojeće metode i tehnike za automatizaciju procesa proizvodnje i poboljšanje proizvoda. U proizvodnji sveže hrane je često potrebno izdvojiti neispravan proizvod od ispravnog na osnovu razlike u boji [1]. Tako, sistem za sortiranje hrane sadrži i optički podsistem koji sadrži kamere za detektovanje neispravnog proizvoda. U mnogim drugim oblastima industrije se koristi robotska ruka kao sastavni deo automatizovanog sistema, a poboljšanje njenog rada može se postići dodavanjem optičkog sistema pomoću kojeg je moguće prepoznati boju ili oblik nekog proizvoda [2]. Ovakvi primeri upućuju na to da i obrazovni sistem u tehničkim disciplinama treba da se razvija u sličnom smeru, kroz realizaciju novih laboratorijskih postavki posvećenih novim metodama i tehnikama za industrijske procese. Zahvaljujući ubrzanom razvoju digitalne elektronike i mikrokontrolerskih platformi, to se može uraditi korišćenjem gotovih elektronskih komponenti i senzora [3]. U ovom radu je opisana jedna laboratorijska postavka za sortiranje predmeta prema boji bazirana na akviziciji slike pomoću USB kamere i personalnog računara.

Rad je baziran na akviziciji i obradi slike primenom programskog paketa LabVEW [4, 5]. Realizovan je sistem koji vrši detekciju boje određenog predmeta pomoću kamere, i na osnovu detektovane boje vrši sortiranje i merenje mase sortiranih predmeta.

Na početku rada su date su osnovne informacije o upravljačkom programu NI Vision Acquisition Software, namenjenog za akviziciju slike sa određenih uređaja, i dodatnog programa NI Vision Development Module, namenjenog za obradu slike u programu LabVIEW. Takođe je objašnjen princip rada funkcije IMAQ ColorLearn koja je korišćena za određivanje boje predmeta.

Zatim je dat prikaz realizovane laboratorijskog postavke kojom se vrši sortiranje predmeta prema detektovanoj boji. Objašnjen je postupak programske kalibracije USB kamere, prikazani su snimci ekrana računara u toku izvršavanja LabVIEW aplikacije u toku kalibracije i merenja, kao i prikaz rezultata merenja. Takođe, data je i odgovarajuća diskusija namene opisanog sistema i njegov značaj u nastavi.

II. LabVIEW programi za akviziciju videa i slike

Za akviziciju videa i slike u programu LabVIEW potrebno je instalirati upravljački program NI Vision Acquisition Software. Nakon instalacije programa u paleti funkcija se formira paleta Vision and Motion unutar koje se nalaze podpalete NI-IMAQ i NI-IMAQdx, sa funkcijama koje se koriste za akviziciju videa i slike sa uređaja priključenih na računar.

Podpaleta NI-IMAQ se koristi za rad sa sledećim uređajima:

- National Instruments Camera Link Frame Grabbers,
- National Instruments Parallel Digital Frame Grabbers,
- National Instruments Analog Frame Grabbers i
- National Instruments 17xx Smart Cameras.

Podpaleta NI-IMAQdx se koristi za rad sa sledećim uređajima:

- Gigabit Ethernet Cameras Supporting GigEVision,
- FireWire IEEE 1394 Cameras,
- USB 2.0 Cameras Supporting Microsoft DirectShow i
- USB 3.0 Cameras Supporting USB3 Vision.

Za akviziciju i obradu videa i slike potrebno je instalirati i dodatni modul NI Vision Development Module. Nakon instalacije modula, u paleti funkcija Vision and Motion se formiraju sledeće podpalete:

- Vision Utilities – koja sadrži podpalete funkcija pomoću kojih se vrši kreiranje memorijskog prostora za sliku, podešavanje ili očitavanje atributa slike, kopiranje slike, definisanje oblasti slike koja se analizira i druge,
- Image Processing – koja sadrži podpalete funkcija

Branko Stojković – Fakultet tehničkih nauka u Čačku, Univerzitet u Kragujevcu, Svetog Save 65, 32000 Čačak, Srbija, (e-mail: stojkovicbranko@mts.rs) – student master studija.

Branko Koprivica – Univerzitet u Kragujevcu, Fakultet tehničkih nauka u Čačku, Svetog Save 65, 32000 Čačak, Srbija, (e-mail: branko.koprivica@ftn.kg.ac.rs).

Alenka Milovanović – Univerzitet u Kragujevcu, Fakultet tehničkih nauka u Čačku, Svetog Save 65, 32000 Čačak, Srbija, (e-mail: alenka.milovanovic@ftn.kg.ac.rs).

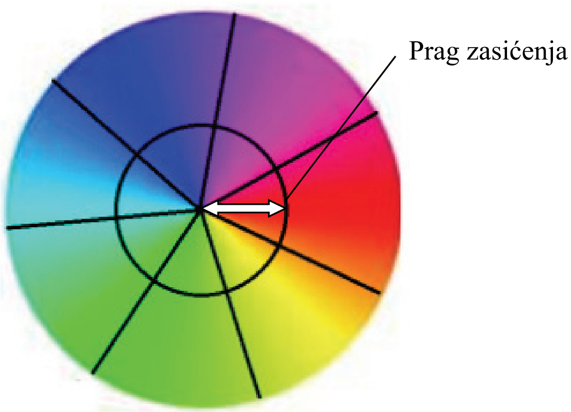
Tatijana Dlabač – Univerzitet Crne Gore, Pomorski fakultet, Dobrota 36, 85331 Kotor, Crna Gora, (e-mail: tanjav@ucg.ac.me).

pomoću kojih se vrši obrada, filtriranje i analiza slike i

- Machine Vision – koja sadrži podpalete funkcija pomoću kojih se vrši merenje rastojanje između detektovanih ivica, detekcija ivica (vertikalne, horizontalne ili kružne), merenje intenziteta piksela, detekcija objekta i druge.

III. Analiza boje slike u LabVIEW programu

Akvizicija videa i slike se najčešće vrši u RGB formatu, ali je za analizu slike pogodniji HSL format [6]. Funkcija IMAQ ColorLearn prvo vrši konverziju slike iz formata RGB u format HSL, a zatim generiše spektar boje (izlaz Color Spectrum) u vidu jednodimenzionalnog niza HSL modela boje od n+2 broja elementa. Poslednja dva elementa niza daju informaciju o beloj i crnoj boji. Broj elemenata u nizu zavisi od izbora nivoa osetljivosti (ulaz Color Sensitivity). Postoji mogućnost izbora tri nivoa osetljivosti nizak (Low), srednji (Medium) i visoki (High). Za nizak nivo osetljivosti, spektar boja HSL modela je podeljen na 7 jednakih sektora, kao što je prikazano na Sl. 1. Za srednji nivo osetljivosti HSL model je podeljen na 14 jednakih sektora, a za visoki nivo osetljivosti na 28 jednakih sektora [6].

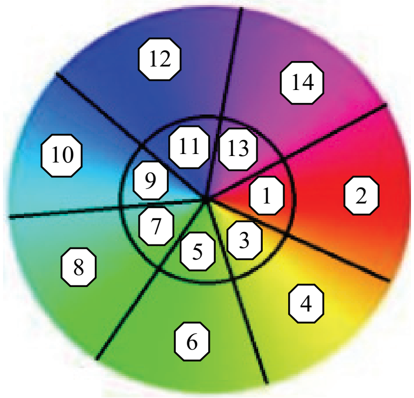


Sl. 1. Ilustracija HSL kruga boja podeljenog na sedam sektora, sa označenim pragom zasićenja.

Ulazom Learn Saturation Threshold funkcije IMAQ ColorLearn se definiše prag zasićenosti neke boje, odnosno granica između svetlije i tamnije nijanse te boje. Vrednost ovog parametra se može podesiti u granicama od 0 do 255 (100 %), gde vrednost 0 odgovara najmanje čistoj nijansi (rezultat je crna boja), a vrednost 255 najčistijoj nijansi (rezultat je detektovana boja). Na ovaj način se svaki sektor može podeliti na dva dela, kao što je to prikazano na Sl. 1. Podrazumevana vrednost je 80 i ona je korišćena za detektovanje boje u ovom radu.

Parametar prag zasićenja deli jedan sektor na dva dela. Tako, broj elemenata izlaznog niza koji opisuje spektar boja iznosi 14 za nizak nivo osetljivosti, kao što je prikazano na Sl. 2, odnosno 28 elemenata za srednji nivo osetljivosti i 56 za visok nivo osetljivosti. Vrednost svakog elementa niza se kreće u granicama od 0 do 1 i predstavlja procenat piksela slike koji se nalazi u nekom od sektora HSL modela boje.

Navedena funkcija ne daje podatak o osvetljenosti analizirane slike.



Sl. 2. Raspored spektra boja u izlaznom nizu za nizak nivo osetljivosti.

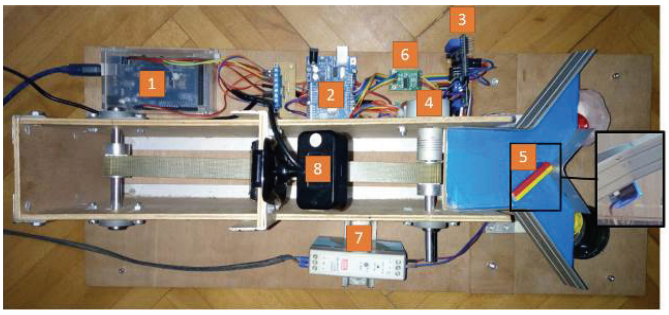
IV. Konstrukcija laboratorijske postavke za sortiranje predmeta prema boji

Realizovana postavka za sortiranje predmeta prema boji i merenje mase sortiranih predmeta, prikazana na Sl. 3, se sastoji od:

1. Arduino MEGA 2560,
2. Arduino UNO,
3. četvorokanalnog drajvera za motore – Arduino L293D,
4. koračnog motora MOTS2 sa pokretnom trakom,
5. servo motora SG90 sa rampom,
6. dve merne ćelije za 1 kg sa pojačavačem HX711,
7. izvora jednosmernog napona MDR-20-12 i
8. USB kamere CNE-CWC1.

Rad koračnog motora koji pokreće traku na koju se postavljaju predmeti kontroliše Arduino UNO kartica (sa upisanim programskim kodom). Rad servo motora koji pomera rampu kontroliše Arduino MEGA kartica i pomoću nje se vrši merenje mase predmeta, pri čemu se podaci o masi predmeta preko serijskog interfejsa prosleđuju ka računaru i LabVIEW aplikaciji, a od računara ka Arduino kartici se prosleđuje podatak o boji predmeta, na osnovu kojeg se zadaje položaj rampe. Rad USB kamere kontroliše samo računar.

Pokretna traka se okreće stalnom brzinom, predmeti se na traku postavljaju ručno, a kamera vrši kontinualno snimanje. Nakon detektovanja boje predmeta, podešava se položaj rampe i vrši sortiranje predmeta koji pada u odgovarajuću posudu. Predviđeno je da se na traci mogu naći predmeti u dve boje, pa se oni i razvrstavaju u dve grupe, smeštanjem u dve posude. Ispod posuda se nalazi merna ćelija pomoću koje se kontinualno vrši merenje mase predmeta.



Sl. 3. Laboratorijska postavka za sortiranje predmeta prema boji i merenje njihove mase - pogled odozgo.

V. Opis LabVIEW aplikacije i postupka izvođenja laboratorijske vežbe

Za potrebe izvođenja laboratorijske vežbe „Sortiranje predmeta prema boji“ je napravljena LabVIEW aplikacija. Ona je podeljena na tri dela po karticama, i to:

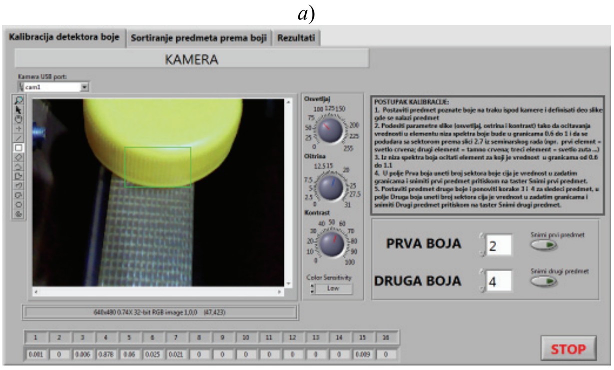
1. Kalibracija detektora boje,
2. Sortiranje predmeta prema boji i
3. Rezultati.

Na početku je potrebno izvršiti kalibraciju detektora boje postavljanjem predmeta poznate boje. Pri tome je najbolje koristiti predmete sa bojom koja je najpribližnija osnovnim bojama. U skladu sa tom preporukom, kalibracija i merenje se može sprovesti na niskom nivou osetljivosti. Predviđeno je da se vrši razvrstavanje dve grupe jednobojnih predmeta, pa je u postupku kalibracije potrebno izvršiti kalibraciju dve boje. Prvo se na traku koja miruje postavi predmet u prvoj boji i podesi se oblast od interesa koja će se koristiti za određivanje boje (zeleni pravougaonik na Sl. 4). Na osnovu vrednosti koje se pokazuju u izlaznom nizu spektra boja se definiše prva boja. Za plastični čep crvene boje je očitana vrednost 0,907 za drugi član niza, dok su svi ostali bili jednaki ili bliske nuli. Shodno tome, prva boja je podešena na vrednost 2, Sl. 4*a*. Postupak se ponavlja sa predmetom u drugoj boji i nakon očitavanja vrednosti u izlaznom nizu se podešava vrednost druge boje. Za žuti plastični čep je očitana vrednost 0,878 u četvrtom članu niza dok su sve ostale vrednosti bile jednake ili bliske nuli, Sl. 4*b*. Druga boja je podešena na vrednost 2.

Ovime se kalibracija detektora boje završava, nakon čega se prelazi na sledeću karticu Sortiranje predmeta prema boji.

Dobijeni rezultati odgovaraju rasporedu spektra boja prikazanom na Sl. 2.

a)



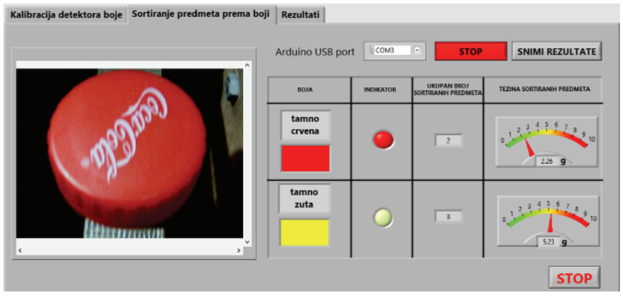
b)

Sl. 4. Izgled LabVIEW aplikacije - kartica Kalibracija detektora boja.

Kartica Sortiranje predmeta prema boji se sastoji od indikatora slike, indikatora boje, numeričkih indikatora za prikaz broja detektovanih predmeta određene boje, grafičkih indikatora za signalizaciju detekcije predmeta određene boje i numeričkih indikatora za ukupnu masu sortiranih predmeta, kao što je prikazano na Sl. 5. Pored indikatora, postavljena je i kontrola za izbor serijskog porta za ostvarivanje komunikacije između računara i Arduino kartica, kao i kontrole za početak očitavanja i snimanje očitanih podataka.

Aplikacija vrši kontinuirano snimanje predmeta koji se kreću po traci i na osnovu podešenih vrednosti boja iz prethodnog koraka vrši se određivanje boje predmeta. Nakon detektovanja boje se menja broj detektovanih predmeta, šalje se komanda motoru sa rampom, u slučaju da je nova boja različita od prethodne, a ujedno se meri i ukupna masa predmeta koji su smešteni u posudama. Omogućeno je ručno snimanje rezultata merenja prema potrebama korisnika.

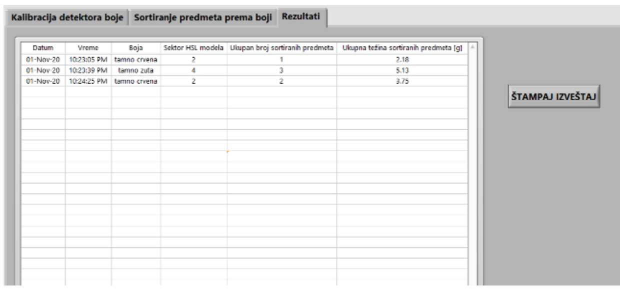
Na Sl. 5 je prikazan izgled ekrana računara u toku izvršavanja ovog dela laboratorijske vežbe u toku kojeg se vrši sortiranje predmeta i merenje njihove mase. Može se uočiti da su uspešno detektovana dva predmeta crvene i tri predmeta žute boje. Izmerene ukupne mase predmeta nakon razvrstavanja su: 2,26 g za crvene čepove i 5,23 g za žute čepove.



Sl. 5. Izgled LabVIEW aplikacije - kartica Sortiranje predmeta prema boji.

Nakon završenog razvrstavanja predmeta se na kartici Rezultati može izvršiti uvid u zabeležene rezultate.

Na Sl. 6 je dat prikaz ove kartice sa zabeleženim rezultatima. U tabeli su dati podaci o datumu i vremenu, detektovanoj boji, sektoru HSL modela za detektovane boje, ukupan broj detektovanih predmeta i njihova ukupna masa.



Sl. 6. Izgled LabVIEW aplikacije - kartica Rezultati.

Nakon uvida u rezultate je moguće izvršiti štampanje izveštaja u PDF formatu. Izveštaj sadrži podatke o kompletnoj proceduri, uključujući kalibraciju boja, detektovanje boja i same rezultate, uključujući i slike predmeta.

VI. Zaključak

Ovaj rad prikazuje realizaciju laboratorijske postavke za sortiranje predmeta prema boji. Sortiranje je zasnovano na akviziciji videa sa USB kamere priključene na računar i obradi slike u programu LabVIEW.

Na osnovu snimljene slike se vrši određivanje spektra boja HSL modela, a na osnovu tog spektra se vrši kalibracija detektora boje i određivanje boje snimljenog predmeta. Prema detektovanoj boji se vrši sortiranje predmeta u dve grupe i vrši se merenje ukupne mase tako sortiranih predmeta.

Rad upotrebljenih komponenti (motora, mernih ćelija i kamere) kontroliše se pomoću Arduino kartica sa programskim kodom i računara sa LabVIEW aplikacijom. Između računara i Arduino kartica je ostvarena serijska komunikacija za razmenu podataka za kontrolu i rezultata merenja mase sortiranih predmeta.

U radu je opisana LabVIEW korisnička aplikacija za upravljanje radom realizovanog sistema i očitavanja rezultata merenja sa senzora. Ona se sastoji iz tri dela: prvi za kalibraciju detektora boje, drugi za razvrstavanje predmeta po boji i merenje njihove mase i treći za prikaz sačuvanih rezultata. Krajnji korak u izvođenju laboratorijske vežbe je kreiranje izveštaja koji sadrži sve relevantne podatke, uključujući i slike detektovanih predmeta.

Takođe, u radu je opisan i konkretan primer razvrstavanja plastičnih čepova crvene i žute boje, dati su rezultati dobijeni u toku kalibracije i u toku merenja i njihova diskusija.

Ovaj rad je nastao u sklopu studentskog istraživanja u okviru predmeta Virtuelna instrumentacija i Električna merenja neelektričnih veličina na master studijama Elektrotehničkog i računarskog inženjerstva na Fakultetu tehničkih nauka u Čačku. To istraživanje je sastavni deo obimnijeg istraživanja koje zajedno sprovode FTN Čačak i Pomorski fakultet Kotor sa ciljem unapređenja praktične nastave. Opisana laboratorijska postavka će biti korišćena u realizaciji praktičnog dela nastave sa budućim generacijama studenata. Na taj način će se studenti upoznati sa osnovnim pojmovima i metodama za akviziciju i obradu videa i slike. Takođe, biće u mogućnosti da istu unaprede kroz hardverska i programska poboljšanja, kao i da na sličnom principu osmisle nove postavke. Poboljšanja ili nova rešenja mogu biti posvećena drugim aspektima obrade slike, određivanja kontura i dimenzija predmeta, poređenju slika po obliku i dimenzijama, prebrojavanju objekata na slici i drugo. Takođe, mogu se realizovati i sistemi sa većom brzinom rada i boljim kvalitetom slike, prema uzoru na neka industrijska rešenja.

Zahvalnica

Ovaj rad je rezultat bilateralnog projekta koji finansiraju Ministarstvo prosvete, nauke i tehnološkog razvoja Republike Srbije i Ministarstvo prosvjete, nauke, sporta i kulture (prethodno Ministarstvo nauke) Crne Gore.

Literatura

[1] J.M. Low, W.S. Maughan, S.C. Bee, M.J. Honeywood, “ Sorting by colour in the food industry,” in *Instrumentation and Sensors for the Food Industry*, ch. *5*, pp. 117–136*,* Cambridge, England/UK: Woodhead Publishing, 2001.

[2] E. Apriaskar, Fahmizal, M. R, Fauzi, “Robotic technology towards industry 4.0: Automatic object sorting robot arm using kinect sensor,” *J. Phys. Conf. Ser.*, vol. *1444,* p. *012030,* Feb., 2020.

[3] M. A. Alaya, Z. Tóth, A. Géczy, “Applied Color Sensor Based Solution for Sorting in Food Industry Processing,” *Period. Polytech. Electr. Eng. Comput. Sci.*, vol. *63,* no. *1,* pp. *16-22,* Feb., 2019.

[4] A. Milovanović, M. Bjekić, B. Koprivica, *Virtuelna instrumentacija*, Čačak, Srbija: UKG-TF Čačak, 2010.

[5] M. Janković, M. Barjaktarović, M. Novičić, P. Atanasijević, *Praktikum iz merno-akvizicionih sistema,* Beograd, Srbija: UB-ETF, 2019.

[6] National Instruments Corporation, „IMAQ Vision Concepts Manual”, Austin, Texas, USA, 2003.

Abstract

The aim of this paper is to present a laboratory setup for sorting of objects according to colour by video acquisition in the LabVIEW software package. The paper gives basic information about the software add-ons required to work with video and images with LabVIEW software and their possibilities for application and briefly describes the HSL image conversion system. The construction of the model for sorting of objects by colour, software calibration of the USB camera and the procedure of the laboratory exercise are also described in the paper. At the end of the paper, the results of measurement of the mass of the sorted objects are presented as well as appropriate discussion.

Sorting of objects according to colour by video acquisition using virtual instrumentation

Branko Stojković, Branko Koprivica, Alenka Milovanović and Tatijana Dlabač

Mogućnost primene Hamonovih presloživih otpornika u naizmeničnom režimu

Stefan Mirković, Dragan Pejić, Marina Subotin, Nemanja Gazivoda, Zdravko Gotovac

***Apstrakt*—U većini primena u praksi Hamonovi (presloživi) otpornici nalaze svoje mesto u DC režimu. Njihova prvenstvena uloga nije da budu etaloni otpornosti, već etaloni prenosnog odnosa otpornosti. U ovom radu su izneti osnovni predlozi primene Hamonovih otpornika u AC režimu. Zbog svojih dobrih metroloških karakteristika u DC režimu, u radu su postavljene hipoteze koje pokušavaju da te karakteristike iskoriste u AC režimu.**

***Ključne reči*—merenje;odnos;etalon;standard;otpornik.**

I. Hamonovi otpornici

Pod Hamonovim otpornicima se smatra grupa od n otpornika iste nazivne vrednosti koji su međusobno trajno redno povezani preko četvorožičnih spojeva. Otpornost izvoda ovih spojeva uračunata je u otpornost pojedinačnih otpornika. Kod Hamonovih otpornika, teži se da se postigne visoka tačnost pojedinačnih otpornika. Na naponske i strujne izvode otpornika dodaju se kompenzacioni otpornici koji svi zajedno čine kompenzacionu mrežu. Prespajanje u paralelnu vezu vrši se preko te specijalne kompenzacione mreže otpornika i pomoću posebnih kratkospojnika. Za n redno vezanih otpornika pri povezivanju u paralelnu vezu postiže se smanjenje otpornosti za veoma približno n^2 puta. Odnos otpornosti redne i paralelne veze naziva se i transferom Hamonovog otpornika. Kako bi se potisnuli negativni efekti otpornosti spojeva u cilju povećanja tačnosti Hamonovog transfera, koriste se četvorožični spojevi otpornika. Grešku transfera izazivaju takozvane prolazne (podužne i poprečne) otpornosti četvorožičnih spojeva. Ovaj negativni efekat se kompenzuje dodavanjem pomenutih kompenzacionih otpornika koji realizuju kompenzacionu mrežu.

Posmatra se grupa od n otpornika iste nazivne vrednosti. Stvarna vrednost pojedinačnog otpornika je:

$$R_i = R + \Delta R_i \tag{1}$$

gde je R srednja vrednost grupe otpronika a ΔR_i razlika između stvarne vrednosti pojedinačnog otpornika i srednje vrednosti grupe. Srednja vrednost grupe otpornika je:

$$R = \frac{1}{n}\sum_{i=1}^{n} R_i \tag{2}$$

Ako se (1) svede na sledeći način:

$$R_i = R\left(1 + \frac{\Delta R_i}{R}\right) \tag{3}$$

može se reći da je

$$R_i = R\left(1 + \delta_{R,i}\right) \tag{4}$$

gde je $\delta_{R,i}$ relativno odstupanje pojedinačnih otpornika od srednje vrednosti grupe. Ekvivalentna otpornost redne veze n otpornika (R_S) je:

$$R_S = \sum_{i=1}^{n} R_i = \sum_{i=1}^{n} R\left(1+\delta_{R,i}\right) = nR\left(1 + \frac{1}{n}\sum_{i=1}^{n}\delta_{R,i}\right) \tag{5}$$

Pošto je zbir odstupanja pojedinačnih uzoraka od aritmetičke sredine nekog skupa jednak nuli, to znači da je član:

$$\sum_{i=1}^{n}\delta_{R,i} = 0 \tag{6}$$

Prema tome, (5) dobija sledeći oblik:

$$R_S = \sum_{i=1}^{n} R_i = nR \tag{7}$$

Ekvivalentna otpornost paralelne veze n otpornika (R_P) može se odrediti na osnovu sledećeg izraza:

$$\frac{1}{R_P} = \sum_{i=1}^{n}\frac{1}{R_i} \Rightarrow R_P = \frac{1}{\sum_{i=1}^{n}\frac{1}{R\left(1+\delta_{R,i}\right)}} = \frac{R}{\sum_{i=1}^{n}\frac{1}{\left(1+\delta_{R,i}\right)}} \tag{8}$$

Odnos ekvivalentne otpornosti serijske i paralelne veze n otpornika je:

$$\frac{R_S}{R_P} = \frac{nR}{\frac{R}{\sum_{i=1}^{n}\frac{1}{\left(1+\delta_{R,i}\right)}}} = n\sum_{i=1}^{n}\frac{1}{\left(1+\delta_{R,i}\right)} \tag{9}$$

Za $|\delta_i| < 1$, može se primeniti razvoj elementarne funkcije u Maklorenov red:

$$\frac{1}{\left(1+\delta_{R,i}\right)} = \sum_{k=0}^{\infty}\left(-\delta_{R,i}\right)^k = 1 - \delta_{R,i} + \delta_{R,i}^2 - \delta_{R,i}^3 + \ldots \tag{10}$$

Zbog (6) dolazi do poništavanja linearnih članova u sledećem izrazu, i uz zanemarivanje članova 3. i višeg reda, dobija se sledeća aproksimacija:

Stefan Mirković – Fakultet tehničkih nauka, Univerzitet u Novom Sadu, Trg D. Obradovića 6, Novi Sad, Srbija (e-mail: mirkovicst@uns.ac.rs).
Dragan Pejić – Fakultet tehničkih nauka, Univerzitet u Novom Sadu, Trg D. Obradovića 6, Novi Sad, Srbija (e-mail: pejicdra@uns.ac.rs).
Marina Subotin – Fakultet tehničkih nauka, Univerzitet u Novom Sadu, Trg D. Obradovića 6, Novi Sad, Srbija (e-mail: marina.bulat@uns.ac.rs).
Nemanja Gazivoda – Fakultet tehničkih nauka, Univerzitet u Novom Sadu, Trg D. Obradovića 6, Novi Sad, Srbija (e-mail: nemanjagazivoda@uns.ac.rs).
Zdravko Gotovac – Fakultet tehničkih nauka, Univerzitet u Novom Sadu, Trg D. Obradovića 6, Novi Sad, Srbija (e-mail: zdravko.gotovac@uns.ac.rs).

$$\sum_{i=1}^{n}\frac{1}{\left(1+\delta_{R,i}\right)}\approx\sum_{i=1}^{n}\left(1-\delta_{R,i}+\delta_{R,i}^{2}\right)=\sum_{i=1}^{n}\left(1+\delta_{R,i}^{2}\right) \quad (11)$$

Na osnovu toga, (9) dobija sledeći oblik:

$$\frac{R_S}{R_P}=n\sum_{i=1}^{n}\frac{1}{\left(1+\delta_{R,i}\right)}\approx n^2\left(1+\frac{1}{n}\sum_{i=1}^{n}\delta_{R,i}^{2}\right) \quad (12)$$

Poslednji izraz pokazuje da je odnos redne i paralelne veze (transfer) istih n otpornika sa velikom tačnošću jednak kvadratu broja otpornika. Ako se pri ovoj analizi zanemare otpornosti spojeva, može se videti da je glavni uzročnik greške transfera rasipanje vrednosti otpornika oko srednje vrednosti grupe. Relativna vrednost ove greške je:

$$\delta_{DC}=\frac{n^2-\dfrac{R_S}{R_P}}{\dfrac{R_S}{R_P}}\approx\frac{-\dfrac{1}{n}\sum\limits_{i=1}^{n}\delta_{R,i}^{2}}{1+\dfrac{1}{n}\sum\limits_{i=1}^{n}\delta_{R,i}^{2}}=-\frac{\left(\dfrac{\sigma(R_i)}{R}\right)^2}{1+\left(\dfrac{\sigma(R_i)}{R}\right)^2} \quad (13)$$

Greška odnosa (transfera) približno zavisi od varijanse otpornosti grupe. Pri merenju odnosa otpornosti (transfera), nije od velike važnosti tačnost pojedinačnih otpornika već njihova međusobna ujednačenost. Takođe nije potrebna ni visok dugoročna stabilnost otpornosti. Potrebno je i dovoljno da odgovarajuća stabilnost bude održana unutar intervala trajanja jednog merenja. Visoka tačnost transfera važi samo za velike otpornosti, kada se otpornosti spojnih vodova mogu zanemariti. Kod otpornosti pojedinačnih otpornika ispod 100 Ω otpornosti spojnih vodova mogu uzrokovati značajne greške, pa se kompenzacione mreže obavezno koriste.

Hamonovi etaloni pored svoje prvenstvene uloge da budu etaloni prenosnog odnosa otpornosti, nalaze primenu i kao etaloni otpornosti. Razlog ovome je što se oni prave od vrlo kvalitetnih otpornika pa se po svojoj tačnosti mogu porediti sa radnim etalonima. Međutim, pažljivom izradom pojedinačnih otpornika i spojeva, postiže se da greška transfera bude znatno manja od grešaka samih otpornika. Kod poređenja dve otpornosti Hamonovim etalonom, jedna otpornost se poredi sa Hamonovim etalonom kojem su otpornici presloženi u rednu vezu, a drugi otpornik sa istim Hamonovim etalonom kojem su otpornici presloženi u paralelnu vezu. Sada kada se raspolaže sa dva različita rezultata poređenja, traženi odnos te dve otpornosti se računa kao količnik ta dva rezultata poređenja pomnožen sa n^2, tj. sa transferom tog Hamonov etalona. Kod komercijalnih Hamonovih etalona greška transfera može biti ispod 1 ppm. Metodom indirektnog poređenja dve otpornosti pomoću odgovarajućeg Hamonovog etalona postiže se visoka tačnost. Na primer, kod prenošenja vrednosti otpornosti reprodukovane kvantnim Holovim (Hall) efektom na vrednost od 1 Ω i pri dvostrukom transferu 10 000 Ω : 100 Ω i 100 Ω : 1 Ω, postignuta je merna nesigurnost krajnjeg rezultata od oko 10^{-8}.

II. Otpornik kao impedansa

Glavna hipoteza ovog rada se odnosi na to da li ono što važi u DC režimu, na neki način može da važi i u AC režimu, kada se umesto otpornika posmatraju impedanse (moduli i faze)?

U cilju procene ponašanja Hamovonih otpornika ako bi se posmatrali kao impedanse, koje imaju svoj moduo i svoju fazu, izvršene su pre svega matematičke procene, a zatim i računarske simulacije. Glavna ideja je bila da se uoči zavisnost transfera grupe otpornika od promene modula i faze pojedinačnih otpornika. Posmatrana je grupa od n impedansi, sa istim nazivnim modulom i fazom. Tolerancija nazivne vrednosti modula za sve impedanse grupe je ista. Isto važi i za fazu.

Ako se umesto grupe otpornika posmatra grupa od n impedansi, stvarna vrednost pojedinačne impedanse je:

$$\underline{Z_i}=R_i+jX_i=Z_i\cdot e^{j\theta_i} \quad (14)$$

gde je Z_i moduo pojedinačne impedanse, a θ_i njena faza. Aritmetička sredina impedansi grupe je:

$$\underline{Z}=\frac{1}{n}\sum_{i=1}^{n}\underline{Z_i}=R+jX=Z\cdot e^{j\theta} \quad (15)$$

Z predstavlja moduo aritmetičke sredine grupe, a θ fazu aritmetičke sredine grupe. Moduo i faza pojedinačne impedanse može se napisati u sledećem obliku:

$$Z_i=Z\left(1+\delta_{Z,i}\right) \quad (16)$$

$$\theta_i=\theta\left(1+\delta_{\theta,i}\right) \quad (17)$$

gde je $\delta_{Z,i}$ relativno odstupanje modula pojedinačne impedanse od modula aritmetičke sredine grupe. Isto tako, $\delta_{\theta,i}$ je relativno odstupanje faze pojedinačne impedanse od faze aritmetičke sredine grupe. Izraz za ekvivalentnu impedansu serijske veze impedansi grupe je:

$$\underline{Z_S}=\sum_{i=1}^{n}\underline{Z_i}=Z_S e^{j\theta_S} \quad (18)$$

Na osnovu (15) i (18) može se napisati da je:

$$\underline{Z_S}=n\underline{Z} \quad (19)$$

Izraz za admitansu paralelne veze impedansi grupe je:

$$\frac{1}{\underline{Z_P}}=\sum_{i=1}^{n}\frac{1}{\underline{Z_i}}=\frac{1}{Z_P e^{j\theta_P}} \quad (20)$$

Pomoću (19) i (20) , dobija se da je odnos impedansi serijsko i paralelno vezanih impedansi grupe:

$$\frac{\underline{Z_S}}{\underline{Z_P}}=\frac{Z_S e^{j\theta_S}}{Z_P e^{j\theta_P}}=n\underline{Z}\sum_{i=1}^{n}\frac{1}{\underline{Z_i}}=\sum_{i=1}^{n}\frac{nZe^{j\theta}}{Z_i e^{j\theta_i}} \quad (21)$$

Uzimanjem u obzir (16) i (17), prethodni izraz se može napisati u sledećem obliku:

$$\frac{\underline{Z_S}}{\underline{Z_P}}=n\sum_{i=1}^{n}\frac{1}{(1+\delta_{Z,i})}e^{-j(\theta\cdot\delta_{\theta,i})} \quad (22)$$

Sređivanjem (22) dobija se:

$$\frac{Z_S}{Z_P}e^{j(\theta_S-\theta_P)}=n\sum_{i=1}^{n}\frac{1}{(1+\delta_{Z,i})}e^{-j(\theta\cdot\delta_{\theta,i})} \quad (23)$$

Prevođenjem leve i desne strane prethodne jednakosti iz polarnog u Kartezijanski (Dekartov) sistem, dobijaju se sledeće dve jednakosti:

$$\frac{Z_S}{Z_P}\cos(\theta_S-\theta_P)=n\sum_{i=1}^{n}\frac{1}{(1+\delta_{Z,i})}\cos(\theta\cdot\delta_{\theta,i}) \qquad (24)$$

$$-\frac{Z_S}{Z_P}\sin(\theta_S-\theta_P)=n\sum_{i=1}^{n}\frac{1}{(1+\delta_{Z,i})}\sin(\theta\cdot\delta_{\theta,i}) \qquad (25)$$

$$(\theta_S-\theta_P)=\arctan\left(-\frac{\sum_{i=1}^{n}\frac{1}{(1+\delta_{Z,i})}\sin(\theta\cdot\delta_{\theta,i})}{\sum_{i=1}^{n}\frac{1}{(1+\delta_{Z,i})}\cos(\theta\cdot\delta_{\theta,i})}\right) \qquad (26)$$

Uvođenjem smene i rešavanjem ove dve jednakosti, dobijaju se krajnji izrazi koji pokazuju kakav je moduo i faza odnosa serijske i paralelne veze impedansi grupe:

$$\frac{Z_S}{Z_P}=\frac{n}{\cos(\theta_S-\theta_P)}\sum_{i=1}^{n}\frac{1}{(1+\delta_{Z,i})}\cos(\theta\cdot\delta_{\theta,i})=$$
$$-\frac{n}{\sin(\theta_S-\theta_P)}\sum_{i=1}^{n}\frac{1}{(1+\delta_{Z,i})}\sin(\theta\cdot\delta_{\theta,i}) \qquad (27)$$

Z_S / Z_P predstavlja moduo navedenog odnosa, a $(\theta_S - \theta_P)$ njegovu fazu (fazu transfera). U (26) i (27) argument funkcija *sin* i *cos* je ugao $\theta\cdot\delta_{\theta,i}$, koji ustvari predstavlja razliku faza pojedinačne impedanse i aritmetičke sredine grupe.

Vidi se da u ovim izrazima vrednost modula pojedinačnih impedansi ne utiče na rezultat, dok njihove tolerancije, odnosno rasipanje tih modula utiče. Što se tiče faze, vidi se da u ovim izrazima figuriše ugao $\theta\cdot\delta_{\theta,i}$. Utvrđeno je da se ugao $\theta\cdot\delta_{\theta,i}$ nalazi u opsegu koji je veoma sličan opsegu definisanim nazivnom vrednošću faze pojedinačnih impedansi i njihovim tolerancijama.

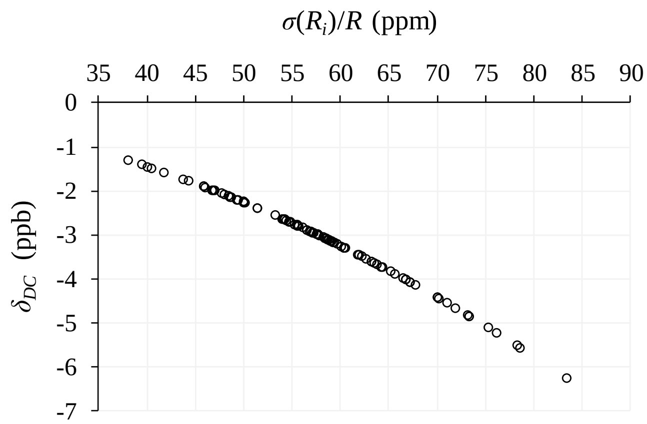
Ako se pogleda (22), izraz za transfer impedansi, i ako se pretpostavi da su sve faze impedansi jednake, bez obzira koliko one iznose, ugao $\theta\cdot\delta_{\theta,i}$ jednak je nuli. Pod ovim uslovom, kompleksni deo u (22) nestaje. Ovo dovodi do sledećeg zaključka da što je rasipanje faza impedansi grupe manje, transfer te grupe je bliži realnom broju. Ako se pretpostavi idealan slučaj da je $\sigma(R_i)$ i $\sigma(X_i)$ jednako nuli, transfer impedansi će biti n^2, bez obzira na to kolika je faza impedansi. Relativna greška transfera može se definisati kao:

$$\underline{\delta_{AC}}=\frac{n^2-\underline{\underline{\frac{Z_S}{Z_P}}}}{\underline{\underline{\frac{Z_S}{Z_P}}}} \qquad (28)$$

III. Rezultati

U cilju dobijanja konkretnih rezultata kao smernica za dalja istraživanja, izvršena je simulacija kojom je određivan odnos serijske i paralelne veze *n* impedansi. Pretpostavka je bila da su simulirane takve impedanse, čiji bi moduli, rasipanje tih modula, faza i rasipanje tih faza odgovaralo realnim situacijama.

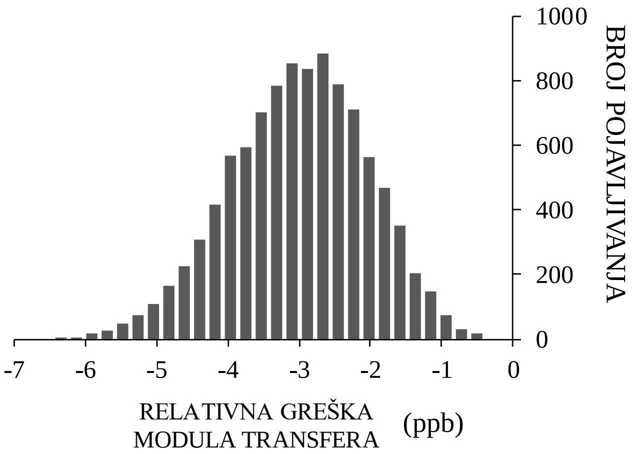
Prvi slučaj je simulacija određivanja transfera *n* otrpornika (DC režim). Simuliranje je vršeno tako što je definisana nazivna vrednost otpornika, kao i tolerancija. Na osnovu tih parametara, izgenerisano je *n* otpornosti. Zatim je određena ekvivalentna serijska, kao i paralelna otpornost. Nakon toga, određen je količnik (odnos) otpornosti serijske i paralelne veze, tj. transfer. Simulirana je situacija sa n=10 otpornika nazivne vrednosti R_n sa tolerancijom 100 ppm prema uniformnoj raspodeli. Određivanje odnosa je ponovljeno 100 puta, gde su dobijana različita rasipanja grupe otpornika. Na sledećoj slici nalazi se zavisnost greške transfera od relativne standardne devijacije (tj. koeficijenta varijacije) grupe otpornika.



Sl. 1. Greška transfera u DC režimu

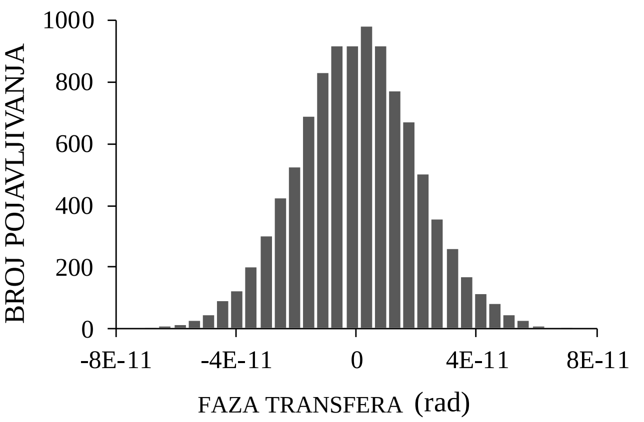
Kako se očekivanja rezultata simulacija zasnivaju na teorijskom modelu (13), dobijeni su očekivani rezultati, i u skladu su sa teorijskim modelom. Ako se posmatra grupa od *n* otpornika nazivne vrednosti R_n sa tolerancijom 100 ppm sa uniformnom raspodelom, matematički (statistički) procenjena relativna standardna devijacija te grupe bila bi približno 58 ppm. Prema (13), grešku transfera treba očekivati da je približno $-(58\text{ ppm})^2/(1+(58\text{ ppm})^2)$=-0,0033 ppm = -3,3 ppb (pod “ppb” se smatra da je to relativna vrednost 10^{-9}, odnosno 1:10^9). Prethodni grafik sa rezultatima simulacija to potvrđuje. Greška transfera približno je jednaka kvadratu koeficijenta varijacije grupe. Za približnu procenu maksimalne greške transfera, uzima se da je koeficijent varijacije grupe otpornika jednak njihovoj toleranciji. Može se videti da je relativna vrednost greške transfera značajno manja od tolerancija samih otpornika kojima se realizuje taj transfer. Upravo zbog ovoga Hamonovi etaloni obezbeđuju visoku tačnost pri prenosu jedne otpornosti na drugu.

Drugi slučaj je simulacija određivanja transfera *n* impedansi (AC režim). Simuliranje je vršeno na isti način kao i prethodno, ali su umesto otpornosti u račun uzete impedanse. Simulirano je n=10 impedansi nazivnog modula Z_n sa tolerancijom od 100 ppm (uniformna raspodela), i nazivne faze θ_n = 0.01 rad sa tolerancijom od 100 ppm (uniformna raspodela). Simulacija je izvršena u 10 000 iteracija.



Sl. 2. Histogram greške modula transfera u AC režimu

Na prethodnoj slici predstavljen je histogram relativne greške modula transfera (n^2-Z_S/Z_P)/(Z_S/Z_P). Može se videti da se moduo transfera prilično simetrično rasipa oko neke vrednosti. Veoma sličan opseg dobija se i u DC režimu. Može se reći da moduo transfera ima vrlo malo rasipanje i odstupanje od n^2 , što je prvo bitno zapažanje. Na sledećoj slici nalazi se histogram faza transfera.



Sl. 3. Histogram faze transfera u AC režimu

Sledeće veoma bitno zapažanje je da se faza transfera rasipa u veoma uskom opsegu. Razlog tome je što su faze serijske i paralalene veze impedansi iste grupe veoma slične. Primera radi, kada bi se tolerancije modula i faza gore pomenutih impedansi grupe postavile na 1 % (uniformna raspodela), faze serijske od faza paralalene veze bi se razlikovale za približno do 90 ppm. U slučaju da su tolerancije kao i u simulaciji 100 ppm, ove faze bi se razlikove do otprilike 10 ppb.

IV. Zaključak

U situacijama kada se Hamonovi otpornici koriste u DC režimu, pokazuje se da greška transfera zavisi od koeficijenta varijacije grupe otpornika. Što je rasipanje otpornosti oko njihove aritmetičke sredine manje, tim je i greška transfera manja. Simulacije su pokazale da i kod AC režima moduo transfera ima takođe malo odstupanje od n^2. Sledeće bitno zapažanje je da se faza transfera rasipa u okolini ili u blizini nule. Treba obratiti pažnju da se, prema ovoj simulaciji i za ovu grupu otpornika, greška transfera u DC režimu i greška modula transfera u AC režimu, nalaze u sličnom opsegu, i reda su ppb-a.

Ove simulacije, kao i prikazane matematičke hipoteze, daju ohrabrujuće rezultate za dalja istraživanja i primenu Hamonovih etalona u AC režimu.

Literatura

[1] “SR-1010 Series Resistance Transfer Standards USer and Service Manual,” *IET LABS, INC.*, 534 Main Street, Westbury, NY 11590, November, 2008.

[2] R. Radetić, *Električna otpornost: pojava i merenja: sa originalnim rešenjima autora*, Agencija Eho, Niš, 2015.

[3] D. G. Jarrett, “Evaluation of Guarded High-Resistance Hamon Transfer Standards,” *IEEE Transactions on Instrumentation and Measurement*, vol. *48*, no. *2*, April, 1999.

[4] B. V. Hamon, “A 1-100 Ω build-up resistor for the calibration of standard resistors”, *Jour. Sci. Instr.*, vol. 31, pp. 450-453, 1954.

[5] J. Bohacek, “Evaluation of frequency performance of resistance standards”, IMTC/2002. Proceedings of the 19th IEEE Instrumentation and Measurement Technology Conference, 2002.

[6] J. C. Riley, “The Accuracy of Series and Parallel Connections of Four-Terminal Resistors”, *IEEE Transactions on Instrumentation and Measurement*,, vol. 16, 1967.

[7] White, D.R., Jones, K., Williams, J.M., Ramsey, I.E., “A simple resistance network for calibrating resistance bridges”, *IEEE Transactions on Instrumentation and Measurement*,, vol. 46, 1997.

[8] L. Cimeanu, M. Simionescu, “Metrological characterization of reference standard resistors group of 100 Ω by means of Hamon resistor,” 2012 International Conference and Exposition on Electrical and Power Engineering, Iasi, Romania, 25-27 October, 2012.

Abstract

In most practical applications, Hamon transfer resistors find their place in DC mode. Their primary role is not to be a standards of resistance, but standards of ratio of restance. In this paper are presented the basic proposals for Hamon standard use in AC mode. Due to their good metrological characteristics in DC mode, paper presents hypotheses that try to use these caharacteristics in AC mode.

Possibility of using Hamon resistors in alternating mode

Stefan Mirković, Dragan Pejić, Marina Bulat, Nemanja Gazivoda

Metoda etaloniranja multifunkcijskog kalibratora za ispitivanje bezbednosti električnih instalacija

Đorđe Novaković, *Student Member, IEEE,* Nemanja Gazivoda, *Member, IEEE,* Dragan Pejić, *Member, IEEE,* Marjan Urekar, *Member, IEEE,* Ivan Gutai, *Member, IEEE,* Zdravko Gotovac

***Apstrakt*—U radu je opisan razvoj i validacija merne metode za etaloniranje vremena reakcije zaštitnih uređaja diferencijalne struje (RCD - Residual current device), kao jedne od funkcija multifunkcijskog kalibratora Time Electronics 5030. Metoda je razvijena sa ciljem kompletiranja sistema mernih metoda kojima se realizuje etaloniranje svih funkcija multifunkcijskog kalibratora za ispitivanje bezbednosti električnih instalacija. U ovu svrhu je razvijen namenski firmver i hardver. Teorijski i eksperimentalno je pokazano da predložena metoda može uspešno da se primenjuje za etaloniranje vremena reakcije RCD uređaja. Merna nesigurnost realizovanog mernog sistema je dvadeset puta manja od merne nesigurnosti zadavanja vremena reakcije RCD uređaja u okviru multifunkcijkog kalibratora.**

***Ključne reči*— RCD; STM32; multifunkcijski kalibrator; etaloniranje; merna nesigurnost**

I. UVOD

Laboratorija za metrologiju Fakulteta tehničkih nauka u Novom Sadu predstavlja akreditovanu metrološku laboratoriju za etaloniranje mernih instrumenata u oblastima električnih veličina, termometrije i vremena i frekvencije.

Merne metode koje se koriste u etaloniranjima se konstantno inoviraju, kako bi pratile korak sa razvojem tehnologije i novim trendovima u naučnoj i industrijskoj oblasti metrologije.

Prenosna merila - analizatori bezbednosnih parametara električnih instalacija poslednjih godina postaju sve složenija i nude mogućnost merenja sve većeg broja veličina. U skladu sa tim je Laboratorija za metrologiju prepoznala potrebu za nabavkom kalibratora koji bi predstavljao zaokruženo rešenje za etaloniranje ovakvih analizatora. Nedostatak adekvatne dokumentacije na temu etaloniranja svih funkcija nabavljenog multifunkcijskog kalibratora Time Electronics 5030 je nadomešten analizom i ispitivanjem različitih metoda etaloniranja. Jedina preostala funkcija, za koju nije uspešno odabrana pogodna metoda etaloniranja je bila funkcija pri kojoj multifunkcijski kalibrator simulira rad zaštitnih uređaja diferencijalne struje, odnosno vreme reakcije za definisanu vrednost struje aktivacije.

Zaštitni uređaji diferencijalne struje su električni uređaji koji isključuju električno kolo kada god se detektuje postojanje razlike između vrednosti struje u faznom i nultom provodniku. Pomenuta nesrazmernost struja je rezultat curenja struje ka uzemljenju preko metalne konstrukcije koja se usled kvara nalazi pod naponom. Svaki RCD uređaj, pre svega karakterišu dva parametra: struja reakcije i vreme reakcije.

Dalja podela se vrši prema tipu diferencijalne struje na:

- Tip AC - ovakvi RCD uređaji reaguju samo pri pojavi neželjene naizmenične struje sinusnog talasnog oblika;
- Tip A - ovakvi RCD uređaji reaguju pri pojavi neželjene naizmenične struje sinusnog talasnog oblika i pri pojavi jednosmerne pulsirajuće struje;
- Tip U - RCD uređaji osetljivi na naizmeničnu i pulsirajuću jednosmernu struju, prilagođeni primeni sa frekventnim pretvaračima;
- Tip B - univerzalno osetljiv RCD uređaj koji će reagovati kod svih tipova neželjenih struja (naizmeničnih, pulsirajućih jednosmernih i konstantnih jednosmernih).

Prema vremenu reakcije RCD uređaji se dele na:

- Standardne - imaju definisano samo maksimalno vreme reakcije;
- Tip S - RCD uređaj sa selektivnim vremenom reakcije (minimalno 40 ms) i sa mogućnošću dodatnog kašnjenja u reakciji;
- Tip G - RCD uređaj sa definisanim minimalnim vremenom reakcije od 10 ms.

Multifunkcijski kalibrator Time Electronics 5030 nudi mogućnost simuliranja rada svih pomenutih tipova RCD uređaja, kako bi se omogućilo etaloniranje analizatora bezbednosnih parametara električnih instalacija, kao deo obima akreditacije Laboratorije.

Usledilo je detaljno ispitivanje kako bi se došlo do odabira odgovarajuće metode etaloniranja i zatim je usledio razvoj

Đorđe Novaković – Fakultet tehničkih nauka, Univerzitet u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: djordjenovakovic@uns.ac.rs).

Nemanja Gazivoda – Fakultet tehničkih nauka, Univerzitet u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: nemanjagazivoda@uns.ac.rs).

Dragan Pejić – Fakultet tehničkih nauka, Univerzitet u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: pejicdra@uns.ac.rs).

Marjan Urekar – Fakultet tehničkih nauka, Univerzitet u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: urekarm@uns.ac.rs).

Ivan Gutai – Fakultet tehničkih nauka, Univerzitet u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: gutai@uns.ac.rs).

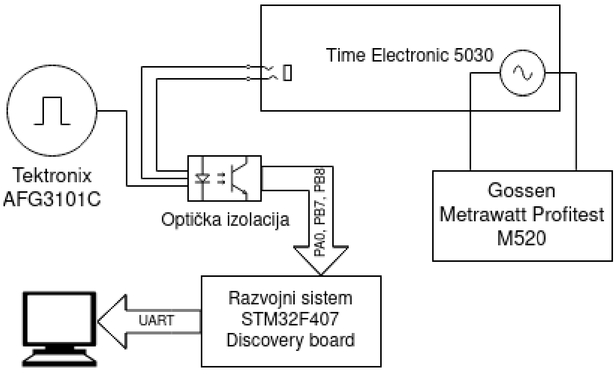
Zdravko Gotovac – Fakultet tehničkih nauka, Univerzitet u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: zdravko.gotovac@uns.ac.rs).

prototipskog sistema za etaloniranje vremena reakcije [1,2].

II. REALIZOVANI MERNI SISTEM - HARDVER

Realizovani merni sistem se sastoji iz kola za izolaciju signala i mikrokontrolerskog kola za akviziciju signala. Izolovani su signali kalibratora Time Electronics 5030, kao i digitalni signal (četvrtke) sa generatora referentnog takta Tektronix AFG3101C. Signali se dovode na mikrokontroler koji vrši merenje vremena na osnovu pristiglih signala. Izmerena vrednost prosleđuje se serijskom portu računara posredstvom standardizovanog RS-232 protokola (UART).

Na slici 1 je prikazana blok šema realizovanog mernog sistema za etaloniranje vremena reakcije RCD uređaja.



Sl. 1 Blok šema mernog sistema

Zbog nedokumentovanog moda kalibracije kalibratora Time Electronics 5030, korišćenjem 3,5 mm stereo audio kontektora posmatrani su signali na izlazu, na zadnjem panelu kalibratora. Pored posmatrana dva signala na izlazu za kalibraciju, posmatran je i signal na šuko utičnici kalibratora koji je snimljen sondom sa faktorom deljenja 1:100. Dobijeni signali prikazani su na slici 2.

Sl. 2 Izgled signala dobijenih akvizicijom sa multifunkcijskog kalibratora Time Electronics 5030

Svi signali su zabeleženi pri funkciji simulacije rada RCD uređaja na multifunkcijskom kalibratoru i nakon aktivacije iste od strane test uređaja-analizatora Gossen Metrawatt Profitest M520.

Signal označen plavom bojom na slici 2 predstavlja trenutak kada sinusni signal prolazi kroz nulu. Ovi pravougaoni impulsi će se pojavljivati sve dok korisnik ne pokrene RCD test na Gossen Metrawatt Profitest M520 test instrumentu. Kada se pokrene RCD test pojavljuje se žuti signal na drugom kanalu audio izlaza kalibratora i traje do trenutka koji je definisao korisnik na kalibratoru. Važno je napomenuti da korisnik može podesiti vrednost sa rezolucijom od 10 ms, što upravo odgovara jednoj poluperiodi (jedan signal detekcije prolaska kroz nulu) mrežnog signala. Eksperimentalno je utvrđeno da kalibrator meri vreme od uzlazne ivice signala plave boje pa sve do silazne ivice signala žute boje. Na slici 3 prikazani su signali zabeleženi pri podešenom parametru vremena reakcije RCD-a od 10 ms, 20 ms, 30 ms i 40 ms.

Sl. 3 Izgled signala dobijenih akvizicijom sa multifunkcijskog kalibratora Time Electronics 5030 pri različitim vrednostima vremena reakcije

III. REALIZOVANI MERNI SISTEM - FIRMVER

Na osnovu opisanog postupka realizovan je i mikrokontrolerski kod za akviziciju signala sa multifunkcijskog kalibratora. Korišćeni mikrokontroler je STM32F407VGT 32 bitni RISC ARM Cortex M4 procesor sa taktom do 168 MHz i sa 1 MB fleš memorije, kao i 196 KB SRAM memorije. Ima mogućnost široke primene zbog velikog broja perifernih jedinica poput: 3 nezavisna 12 bitna AD konvertora, 2 nezavisna 12 bitna DA konvertora, 17 tajmersko brojačkih modula, 3 I2C interfejsa, 6 UART interfejsa, 2 CAN interfejsa itd. [3].

Ovaj kontroler je sastavni deo razvojnog sistema STM32F407 Discovery Board koji je proizveden od strane STMicroelectronics. Razvojni sistem sadrži još i STLink (in-circuit debugger/programmer) za programiranje kontrolera i otklanjanje grešaka, mikrofon i audio pojačavač u klasi D za obradu i reprodukciju zvuka, akcelerometar itd. Jedna od mana ovog razvojnog sistema je nepostojanje FTDI čipa za konverziju sa UART na USB, pa je zbog toga ovaj sistem povezan na STM32F4 Discovery Shield proizveden od strane Mikroelektronike. Pomenuta dodatna ploča pored toga što omogućuje jednostavno povezivanje velikog broja senzora korišćenjem click adaptera, sadrži još FTDI čip. Podešavanje i prilagođavanje pinova, kao i programiranje realizovano je u razvojnom okruženju STM32Cube IDE.

U ovom projektu korišćen je tajmer 5 u modu 32-bitnog brojača spoljašnjih impulsa. Signal koji broji tajmer 5 zapravo predstavlja referentni takt generisan referentnim generatorom takta Tektronix AFG3101C i doveden je na ulaz TIM5_CH1 tajmera, tačnije PA0 pin. Signali koji se dovode sa izlaza kalibratora dovode se na pinove GPIO_EXTI7 i GPIO_EXTI8, tačnije pinove PB7 i PB8, respektivno, a pinovi su podešeni u mod detekcije spoljašnjeg prekida. Takođe, podešen je UART na pinovima PA2 i PA3, sa brzinom prenosa 115 200 b/s, u modu sa 8 bita podataka, jednim stop bitom i bez bita pariteta. Signal koji se dovodi na TIM5_CH1 ulaz tajmera se sinhronizuje sa taktom periferije koja kontroliše rad tajmersko brojačkog modula. Izvor takta je spoljašnji oscilator frekvencije takta 8 MHz koji je integrisan na razvojnom sistemu.

Signal sa spoljašnjeg oscilatora se dovodi na PLL strukturu gde se takt procesora podiže na 168 MHz. Pri ovoj brzini takta procesora, maksimalna frekvencija takta za kontrolu tajmersko brojačkog modula je 84 MHz.

Uzimajući u obzir da je frekvencija referentog takta dovedenog na brojač 100 kHz, kašnjenje signala ne igra bitnu ulogu pa se zato može zanemariti. Takođe, jedan od potencijalnih problema je kašnjenje koje se može desiti u slučaju detekcije prekida i skoka u prekidnu rutinu. U dokumentaciji je navedeno da je kašnjenje procesora od trenutka prekida do prve instrukcije u prekidnoj rutini 12 ciklusa (za odabranu frekvenciju oscilatora kašnjenje je 71,4 ns). Pošto je perioda referentnog signala 10 µs, realizovano je softversko upravljanje brojačkim registrom. Važno je napomenuti da kašnjenje od 12 ciklusa se dešava korišćenjem "bare metal" programiranja [4].

U mernom sistemu kod je realizovan korišćenjem HAL biblioteka [5] (prilikom detekcije prekida poziva se nekoliko wrapper funkcija), što će dodati još par taktova kašnjenja prekidne rutine ali će i dalje perioda referentnog signala biti daleko manja.

Upravljanje bi bilo moguće rešiti hardverskim putem, korišćenjem gejtovanog tajmera, ali za ovaj mikrokontroler postoji mogućnost rada samo sa jednim upravljačkim signalom.

Pošto je realizaciji mernog sistema bilo neophodno meriti vreme od uzlazne ivice jednog do silazne ivice drugog signala upravljanje je rešeno dodavanjem spoljašnjih prekida na pinove PB7 i PB8, uz detekciju odgovarajuće ivice. Takođe, ovaj problem je mogao biti rešen dodavanjem dodatnih hardverskih komponenti, ali zbog već pomenute analize vremenske fleksibilnosti, softversko rešenje se nameće kao jeftinije. Ključna komponenta firmvera je prekidna rutina koja detektuje ivice signala i vrši upravljanje tajmerom. Kod prekidne rutine dat je u poglavlju DODATAK.

U skladu sa oznakama na slici 1, na pinu PB8 mikrokontrolera dovedeni su impulsi detektovanja prolaska signala kroz nulu i označeni su plavom bojom. Pošto se ovi impulsi ponavljaju, korisniku je od važnosti samo poslednji impuls, tačnije neophodno je meriti broj impulsa sa referentnog generatora funkcija od uzlazne ivice signala plave boje do silazne ivice signala žute boje. Kada kontroler detektuje uzlaznu ivicu impulsa, označenog plavom bojom, on resetuje tajmer 5, pa počinje brojanje takta sa referentnog generatora od 0.

Prilikom detekcije silazne ivice signala žute boje tajmer se zaustavlja, njegova vrednost se očitava, konvertuje u string i dalje prosleđuje UART-om.

Priprema slanja i slanje nije neophodno raditi iz prekidne rutine, ali pošto korisnik ne može fizički pokrenuti dva uzastopna testa velikom brzinom ovaj postupak neće uticati na rezultat merenja.

IV. Rezultati merenja i ocena merne nesigurnosti

Na osnovu opisa realizovanog hardvera i firmvera autori su se odlučili za primenu standardne metode poređenja pokazivanja razvijenog etalonskog sistema i ispitivanog uređaja (DUT - device under test) [6].

Dakle poređeno je pokazivanje razvijenog mernog sistema i podatak o postavljenoj vrednosti vremena reakcije simuliranog RCD uređaja na multifunkcijskom kalibratoru Time Electronics 5030.

U svrhu validacije realizovanog mernog sistema i odabrane metode etaloniranja sprovedena je serija od po 10 ponovljenih merenja za 6 različitih vremena reakcije (10 ms, 50 ms, 100 ms, 500 ms, 1000 ms i 2000 ms) i za istu vrednost struje aktivacije od 30 mA.

U tabeli 1 je dat prikaz rezultata merenja za vrednost vremena reakcije od 10 ms.

Tabela 1
Prikaz rezultata merenja za vrednost vremena aktivacije RCD uređaja od 10 ms

Parametri aktivacije RCD uređaja		Rezultati merenja	Aritmetička sredina	Standardna devijacija
Struja	Vreme			
(mA)	(ms)	(ms)	(ms)	(ms)
30	10	10,06	10,085	0,014
		10,09		
		10,09		
		10,09		
		10,06		
		10,09		
		10,09		
		10,09		
		10,09		
		10,10		

Kolona "Vreme" se odnosi na parametar vremena reakcije postavljen na objektu etaloniranja - multifunkcijskom kalibratoru, a kolona "Rezultati merenja" na 10 ponovljenih merenja etalonskim sistemom.

Autori su u svrhu validacije pristupili oceni merne nesigurnosti merenja vremena reakcije razvijenim etalonskim sistemom i poređenjem sa mernom nesigurnošću postavljanja vremena reakcije na multifunkcijskom kalibratoru [7].

Kao zadovoljavajući ishod validacije definisan je zahtev da merna nesigurnost etaloniranja vremena reakcije mora biti barem 5 puta manja od merne nesigurnosti postavljanja vremena reakcije na kalibratoru Time Electronics 5030 [8].

U literaturi [2] je identifikovan podatak o tačnosti postavljanja vremena reakcije RCD uređaja na kalibratoru Time Electronics 5030 od ±0,5 ms.

Greška kalibratora je slučajna veličina, karakterisana pravougaonom uniformnom raspodelom, sa očekivanom vrednosti jednakom nuli i granicama greške ±0,5 ms. Merna nesigurnost tipa B je tada jednaka $0{,}5/\sqrt{3}$ ms. Proširena merna nesigurnost je jednaka $2\cdot(0{,}5/\sqrt{3})$ ms, odnosno ±577 µs.

Matematički model rezultata etaloniranja kalibratora Time Electronics 5030 (DUT) je definisan kao:

$$g = t_x - t_s, \quad (1)$$

$$g = t_x - \left(\frac{n+\Delta n}{f_s+\Delta f_s}\right), \quad (2)$$

gde je:

t_s - Vremenski interval, izmeren razvijenim etalonskim sistemom;

g - Greška etaloniranja DUT;

t_x - Vremenski interval, postavljen na DUT - konstanta;

n - Sadržaj brojača. Slučajna veličina, karakterisana normalnom raspodelom, sa očekivanom vrednosti jednakoj aritmetičkoj sredini od k odmeraka i standardnom devijacijom $s_n/\sqrt{k}$ *, gde je* s_n standardna devijacija odmeraka. Standardna nesigurnost tipa A;

Δn - Korekcija, zbog inherentne greške brojanja. Slučajna veličina, karakterisana pravougaonom uniformnom raspodelom, sa očekivanom vrednosti jednakom nuli i granicama greške ±1. Standardna nesigurnost, tip B, jednaka je $1/\sqrt{3}$;

f_s - Frekvencija etalonskog generatora frekvencije - konstanta;

Δf_s - Korekcija, zbog netačnosti frekvencije etalonskog generatora frekvencije. Slučajna veličina karakterisana pravougaonom uniformnom raspodelom sa očekivanom vrednosti jednakoj korekciji frekvencije (iz sertifikata o etaloniranju) i granicama određenim specifikacijom generatora frekvencije (na primer, granicama greške gg_{f_s}). Merna nesigurnost tipa B.

Standardne merne nesigurnosti $u(n)$, $u(\Delta n)$ i $u(\Delta f_s)$ su definisane kao:

$$u(n) = \frac{s_n}{\sqrt{k}},\ u(\Delta n) = \frac{1}{\sqrt{3}},\ u(\Delta f_s) = \frac{gg_{f_s}}{\sqrt{3}}. \quad (3)$$

Kombinovana merna nesigurnost $u(g)$ je jednaka:

$$u(g) = \left[c_n{}^2\cdot u^2(n) + c_{\Delta n}{}^2\cdot u^2(\Delta n) + c_{\Delta f_s}{}^2\cdot u^2(\Delta f_s)\right]^{1/2}, \quad (4)$$

gde su c_n, $c_{\Delta n}$ i $c_{\Delta f_s}$ koeficijenti osetljivosti definisani kao:

$$\begin{aligned} c_n &= \frac{\partial g}{\partial n} = \frac{1}{f_s+\Delta f_s}, \\ c_{\Delta n} &= \frac{\partial g}{\partial \Delta n} = \frac{1}{f_s+\Delta f_s}, \\ c_{\Delta f_s} &= \frac{\partial g}{\partial \Delta f_s} = -\frac{n+\Delta n}{(f_s+\Delta f_s)^2}. \end{aligned} \quad (5)$$

Nakon zamene (5) u (4) izraz (4) postaje:

$$u(g) = \begin{bmatrix} (\frac{1}{f_s+\Delta f_s})^2\cdot(\frac{s_n}{\sqrt{k}})^2 + (\frac{1}{f_s+\Delta f_s})^2\cdot(\frac{1}{\sqrt{3}})^2 + \\ +(-\frac{n+\Delta n}{(f_s+\Delta f_s)^2})^2\cdot(\frac{gg_{f_s}}{\sqrt{3}})^2 \end{bmatrix}^{1/2}. \quad (6)$$

Uzimajući u obzir da su matematička očekivanja Δn i Δf_s jednaka 0, izraz (6) postaje:

$$u(g) = \begin{bmatrix} (\frac{1}{f_s})^2\cdot(\frac{s_n}{\sqrt{k}})^2 + (\frac{1}{f_s})^2\cdot(\frac{1}{\sqrt{3}})^2 + \\ +(-\frac{n}{(f_s)^2})^2\cdot(\frac{gg_{f_s}}{\sqrt{3}})^2 \end{bmatrix}^{1/2}. \quad (7)$$

Nakon adaptacije izraza (7) dobija se konačan izraz za kombinovanu mernu nesigurnost:

$$u(g) = \left[\left(\frac{1}{f_s}\frac{s_n}{\sqrt{k}}\right)^2 + \left(\frac{1}{f_s}\frac{1}{\sqrt{3}}\right)^2 + \left(\frac{\overline{n}}{f_s^2}\frac{gg_{fs}}{\sqrt{3}}\right)^2\right]^{1/2}. \quad (8)$$

Proširena merna nesigurnost je jednaka:

$$U(g) = k\cdot u(g) = 2\cdot u(g). \quad (9)$$

U nastavku je data tabela 2 sa budžetom merne nesigurnosti za postavljene parametre vremena reakcije od 10 ms i vrednosti struje aktivacije od 30 mA na objektu etaloniranja - multifunkcijskom kalibratoru Time Electronics

5030.

TABELA 2
BUDŽET MERNE NESIGURNOSTI ZA VREME REAKCIJE OD 10 ms

Simbol	Matematičko očekivanje	Standardna nesigurnost	Tip Merne nesigurnosti	Raspodela	Koeficijenti osetljivosti	Doprinos mernoj nesigurnosti
n	1008,5	0,428	A	Normalna	10,0E-6	4,3E-6
Δn	0	0,577	B	Uniformna (pravougaona)	10,0E-6	5,8E-6
f_s	100,0E+3					
Δf_s	0	0,577	B	Uniformna (pravougaona)	100,9E-9	58,2E-9
t_s	10,085E-3					7,19E-6
Standardna merna nesigurnost veličine t_s:				7,2	μs	
Proširena merna nesigurnost veličine t_s:				14,4	μs	

Vremenski interval t_s izmeren etalonskim mernim sistemom za postavljeno vreme reakcije od 10 ms je jednak:

$$t_s = \left(\frac{n + \Delta n}{f_s + \Delta f_s} \right), \tag{10}$$

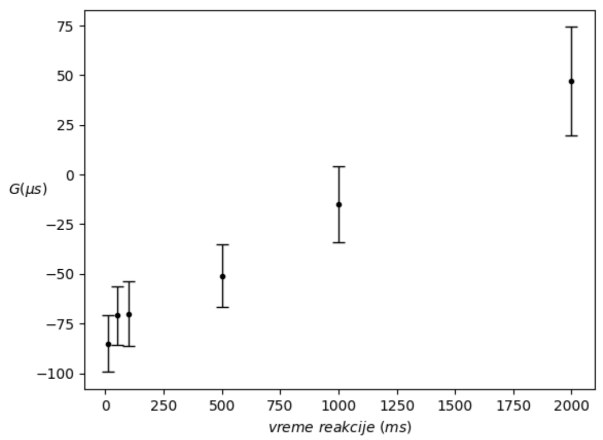
$$t_s = (10,085 \pm 0,0144)\ \text{ms}. \tag{11}$$

U tabeli 3 su prikazani parametri vremena reakcije postavljeni na objektu etaloniranja, greška merenja (razlika između vrednosti postavljene na objektu etaloniranja i vrednosti etalonskog sistema) i vrednosti odgovarajućih mernih nesigurnosti etaloniranja etalonskim mernim sistemom.

TABELA 3
PARAMETRI ETALONIRANJA VREMENA REAKCIJE

Postavljena vrednost vremena reakcije na TE 5030	Greška merenja vremena reakcije tx - ts	Merna nesigurnost merenja vremena reakcije etalonskim mernim sistemom
(ms)	(μs)	(μs)
10	-85	14,4
50	-71	14,8
100	-70	16,1
500	-51	15,8
1000	-15	19,1
2000	47	27,3

Na slici 4 je dat grafik na kojem je prikazana greška merenja vremena reakcije sa pripadajućim vrednostima mernih nesigurnosti u funkciji od 6 različitih vrednosti vremena reakcije.



Sl. 4 Prikaz greške merenja vremena reakcije u funkciji 6 različitih postavljenih vrednosti vremena reakcije na kalibratoru Time Electronics 5030

V. ZAKLJUČAK

U radu je opisana realizovana merna metoda, hardver i firmver kao rešenje problema definisanog u uvodnom poglavlju.

U cilju validacije su izvršena merenja i ocenjena je merna nesigurnost etaloniranja vremenske reakcije Multifunkcijskog kalibratora Time Electronics 5030 u funkciji simulacije RCD uređaja.

Dobijeni rezultati potvrđuju zadovoljavajući ishod validacije primenjene metode etaloniranja uz dobijenu mernu nesigurnost etaloniranja vremena reakcije koja je 20 puta manja od merne nesigurnosti postavljanja vremena reakcije na kalibratoru Time Electronics 5030.

Ovim je potvrđena opravdanost predložene metode etaloniranja i upotrebe realizovanog hardvera i firmvera.

DODATAK

Kod prekidne rutine realizovanog firmvera koja detektuje ivice signala i vrši upravljanje tajmerom je dat u nastavku.

```
void HAL_GPIO_EXTI_Callback (uint16_t GPIO_Pin)
  {
  if (GPIO_Pin == GPIO_PIN_7)
     {
      HAL_TIM_Base_Stop(&htim5) ;
      cnt = __HAL_TIM_GET_COUNTER(&htim5);
      __HAL_TIM_SET_COUNTER(&htim5, 0);
      HAL_TIM_Base_Start(&htim5);
      memset( txt, 0, sizeof(txt));
      itoa(cnt, txt, 10);
      txt[strlen(txt)]= '\n';
      HAL_UART_Transmit(&huart2 , (uint8_t *) txt ,
                                strlen(txt), 1) ;
      }
  if (GPIO_Pin == GPIO_PIN_8)
     {
      HAL_TIM_Base_Stop(&htim5);
      HAL_TIM_SET_COUNTER(&htim5 , 0);
      HAL_TIM_Base_Start(&htim5);
     }
  }
```

ZAHVALNICA

Ovaj rad je podržan od strane Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije u sklopu projekta Inovativna naučna i umetnička isptivanja iz domena delatnosti FTN-a, broj: 451-03-9/2021-14/200156.

LITERATURA

[1] S. Yang and Y. Chau, *Calibration of residual current device (RCD) testers*, NCSLI Measure, vol. 12, no. 4, pp. 28–37, 2018.

[2] Time Electronics Ltd, *5030 Electrical Tester Calibrator*, User manual, July 2013.

[3] STMicroelectronics, *Rm0090 reference manual - Rev.19*, Dostupno:https://www.st.com/resource/en/referencemanual/dm00031020-stm32f405-415-stm32f407-417-stm32f427-437-and-stm32f429-439-advanced-arm-based-32-bit-mcus-stmicroelectronics.pdf

[4] C. Noviello, *Mastering STM32*, Leadpub, 2017

[5] STMicroelectronics, *Description of STM32F4 HAL and low-layer drivers-rev.7*, Dostupno:https://www.st.com/resource/en/usermanual/dm00105879-description-of-stm32f4-hal-and-ll-drivers-stmicroelectronics.pdf, (2021)

[6] JCGM,*Guide to the expression of uncertainty in measurement - Part 6:Developing and using measurement models*, First edition, 2020.

[7] S. L. Ellison and A. Williams, *Quantifying uncertainty in analytical measurement*, 2012.

[8] JCGM., *Evaluation of measurement data - guide to the expression of uncertainty in measurement*, Int. Organ. Stand. Geneva ISBN, vol. 50,p. 134, 2008.

ABSTRACT

This paper describes the development and validation of a measurement method for calibration of the reaction time of residual current devices (RCD), as one of the functions of the Multifunctional Electrical Tester Calibrator Time Electronics 5030. The method was developed to complete a system of measurement methods that are used for calibration of all multifunction calibrator functions for testing the safety of electrical installations. Dedicated firmware and hardware have been developed for this purpose. It has been shown theoretically and experimentally that proposed method can be successfully applied to calibrate the reaction time of an RCD device. The measurement uncertainty of the realized measuring system is twenty times smaller than the measurement uncertainty of setting the reaction time of the RCD within the multifunction calibrator.

Method for Calibration of Multifunction Electrical Tester Calibrator

Đorđe Novaković, Nemanja Gazivoda, Dragan Pejić, Marjan Urekar, Ivan Gutai, Zdravko Gotovac

Automatizacija etaloniranja digitalnih multimetara

Branislav Lukić, Đorđe Novaković, Nemanja Gazivoda, Platon Sovilj, IEEE member

***Apstrakt*—U radu je opisan detaljan postupak realizacije jednog od rešenja automatizacije procesa etaloniranja digitalnih multimetara pomoću kalibratora Times Electronis 5025. Za realizaciju aplikacije za etaloniranje digitalnih multimetara korišten je programski jezik Pajton sa Qt frejmvorkom.**

***Ključne reči---* *automatizacija,* etaloniranje, Pajton, Qt designer.**

I. Uvod

U ovom radu je predstavljeno jedno rešenje automatizacije procesa etaloniranja digitalnih multimetara pomoću aplikacije razvijene u programskom jeziku Pajton.

Oprema kojoj je potreba kalibracija se šalje u metrološku laboratoriju, gde će je kvalifikovani tehničar prilagoditi specifikacijama ili potvrditi da ih već ispunjava, koristeći merne/ispitne instrumente koji i sami moraju ispunjavati stroge zahteve za kalibraciju. Većina komponenata koje se koriste u industriji se može kalibrisati.

Uzimajući u obzir značajno velike cene trenutno dostupnih aplikacija na tržištu za kalibraciju i upravljanje kalibratorom, motivacija za izradu ovog rada je bila da se napravi finansijski pristupačna aplikacija koja će zadovoljavati određene standarde koji su potrebni.

Delovi od kojih se projekat sastoji su:

A. Kalibrator Times Electronics 5025

B. Blok šema

C. Pajton biblioteke

D. Algoritam aplikacije

E. Izgled aplikacije

F. Rezultati merenja

Branislav Lukić – Fakultet Tehničkih Nauka, Novi Sad, Srbija (e-mail: lukic.branislav97@gmail.com).

Đorđe Novaković – Fakultet Tehničkih Nauka, Novi Sad, Srbija (e-mail: djordjenovakovic@ uns.ac.rs).

Nemanja Gazivoda – Fakultet Tehničkih Nauka, Novi Sad, Srbija (e-mail: nemanjagazivoda@ uns.ac.rs).

Platon Sovilj – Fakultet Tehničkih Nauka, Novi Sad, Srbija (e-mail: platon@ uns.ac.rs).

II. Tehničko rešenje

A. Kalibrator Times Electronics 5025

Times Electronics 5025 je precizni kalibrator dizajniran za široku upotrebu. Ovaj kalibrator može da kalibriše razne instrumente i uređaje kao što su merila frekvencije, ommetre, AC-DC voltemetre, termoparove, PT sonde, strujna klešta i mnoge druge merne uređaje.

Dodavanjem posebnih opcija, možemo kalibrisati osciloskope, tajmere/brojače, vatmetre i još mnogo toga.

Slika1. Times Electronics 5025 multi-function calibrator

Funkcijama i opsezima kalibratora se upravlja sa prednje ploče. Na prednjoj ploči se nalaze tasteri za izbor funkcija i podešavanje opsega veličina. Sve ove informacije su prikazane na LED displeju.

Slika2. Ručno podešavanje izlaza kalibratora

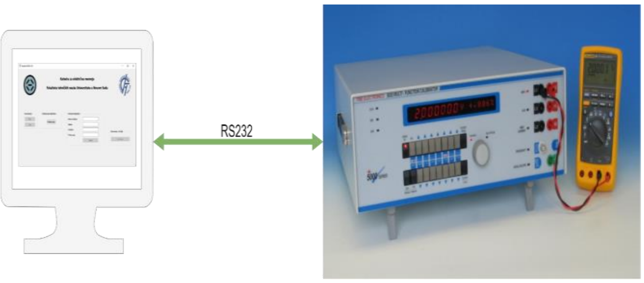
Kalibrator 5025 pored lokalnog upravljanja, ima mogućnost upravljanja preko računara posredstvom RS232 komunikacionog protokola.

Komade za upravljanjem kalibratorom prate SCPI (*Standard Commands for Programmable Instruments*) standard. SCPI standard definiše sintaksu i naredbe koje se koriste za upravljanje programabilnih uređaja za ispitivanje i merenje, kao što su oprema za automatsko ispitivanje i elektronska oprema za testiranje. Preko SCPI komandi se postavljaju opsezi instrumenta, uključuju izlazi itd.

Pored ovih komandi, postoje komande za upite. Pomoću komadi za upit se prikupljaju informacije sa instrumenta. Neke komande mogu istovremeno da postave neku vrednost i postave upit instrumentu.

Sve ove mogućnosti koje pruža kalibrator 5025, dovele su do razvoja aplikacije za automatizaciju procesa etaloniranja digitalnih multimetara.

B. *Blok šema*



Slika3. Blok šema sistema

Na slici 3 je prikazana blok šema celokupnog sistema.

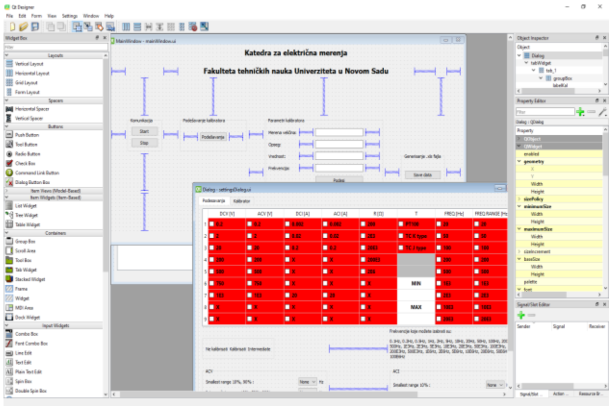
Sistem se sastoji iz dva dela. Prvi deo je računar sa aplikacijom za kalibrisanje sa operativnim sistemom Windows ili Linux, a drugi deo je kalibrator sa multimetrom koji se etalonira.

Ova dva dela su povezana pomoću RS232 komunikacionog protokola. DB9 konektor koji se nalazi na kalibratoru je pomoću adaptera prilagođen na USB radi uspostavljanja komunikacije sa računarom.

C. *Pajton biblioteke*

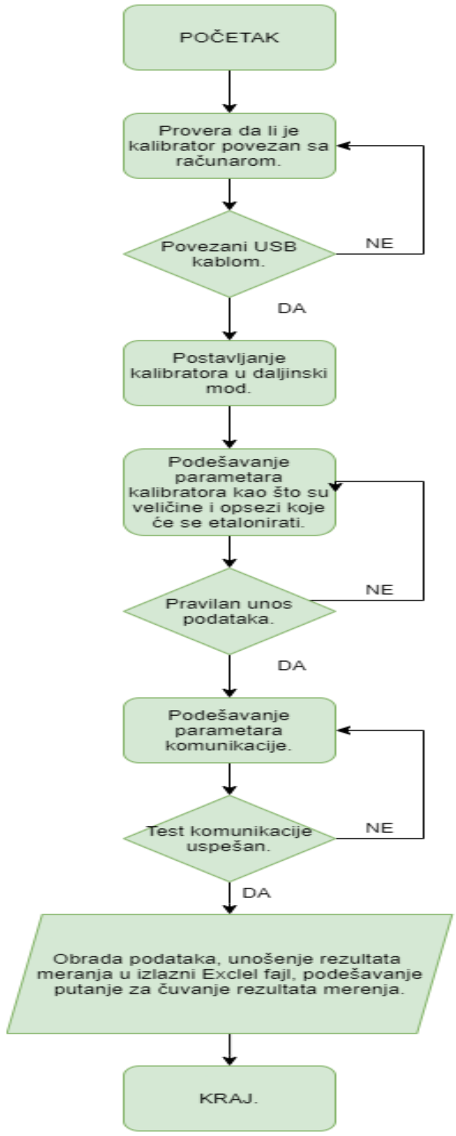
Qt je besplatan alat za pravljenje grafičkih korisničkih insterfejsa (GUI) koji mogu raditi na raznim operativnim sistemima kao što su Windows, Linux, macOS i Android.

PyQt5 je biblioteka koja omogućava korištenje Qt GUI frejmvorka i Pajtona. Sam Qt je napisan u C++ programskom jeziku. Koristeći ga iz Pajtona, nudi nam mogućnost znatne uštede vremena u izradi aplikacije, a da pri tome ne izgubimo brzinu koji nam omogućava C++ .



Slika4. Izleg glavnog prozora QtDesigner-a

D. *Algoritam aplikacije*



Slika5. Algoritam podešavanja tačaka merenja i parametara komunikacije

Prvi korak algoritma se sastoji od provre povezanosti kalibratora I računara USB kablom. Ukoliko provera uspešno prođe prelazi se na drugi korak. Drugi korak algoritma postavlja kalibrator u daljinski mod. Treći korak algoritma se satoji od podešavanja parametara kalibratora kao što su merene veličine, opsezi, i provere ispravnosti unesenih podataka. Ako provera ispravnosti podataka prođe bez greške, dolazi se do drugog koraka gde se podešavaju parametri komunikacije. Ako je komunikacija uspešno uspostavljena, dolazi se do koraka za obradu podataka gde se generiše izlazni fajl sa rezultatima merenja.

Nakon izvršenih svih merenja dolazi se do kraja algoritma.

E. Izgled aplikacije (GUI)

Glavni prozor

Slika6. Prikaz glavnog prozora aplikacije

Glavni prozor se satoji od četiri grupe opcija.

Prvu grupu čini taster "Podešavanje" koje služi za podešavanje parametara kalibratora i generisanje mernih tačaka za izabrane veličine koje će se kalibrisati.

Drugu grupu predstavljaju tastere za početak (*„Start"*) i prekidanje („Stop") komunikacije aplikacije sa kalibratorom 5025. Na taster „Start" i „Stop" nije moguće kliknutu pre nego što se podese parametri komunikacije kao što su port, brzina prenosa, bitovi podataka, stop bit, bit parnosti.

Treću grupu čine parametri kalibratora i taster „Podesi". Parametre kalibratora predstavljaju:

- Merena veličina: DCV, ACV, DCI, ACI, RES, TEMP, FREQ
- Opseg
- Vrednost: ±10%, ±50%, ±90% od opsega
- Frekvencija

U četvrtu grupu opcija spada taster „Save data" koje klikom na njega generiše Excel fajl sa izmerenim podacima koje je operater unosio tokom merenja.

Prozor za podešavanje

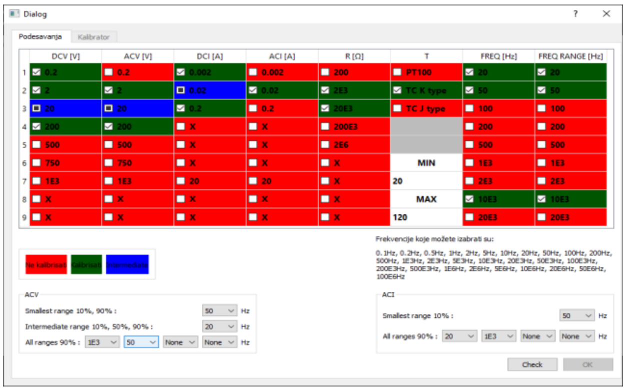
Klikom na taster „Podešavanje" koje se nalazi na glavnom prozoru otvara se novi prozor (slika 7) iz koga se biraju veličine koje će se kalibrisati kao i postavljanje parametara potrebnih za uspostavljanje komunikacije aplikacije sa kalibratorom.

Desnim klikom na polje u tabeli se odabiraju opsezi na kojima će se vršiti kalibracija. Boja polja u tabeli predstavlja informaciju da li se na tom opsegu vrši kalibracija. Crvena boja znači da taj opseg nije odabran, zelena boja daje informaciju da je opseg odabran i plava boja predstavlja srednji opseg.

Pored opsega, za veličine naizmenični napon i naizmeničnu struju se moraju izabrati frekvencije na kojima će se vršiti kalibracija.

Kada se kalibriše senzor za temperaturu, mora se uneti minimalna i maksimalna temperatura senzora koju on meri.

Na slici 7 prikazan je prozor za podešavanje parametara u kome su izabrane vrednosti koje će biti kalibrisane.



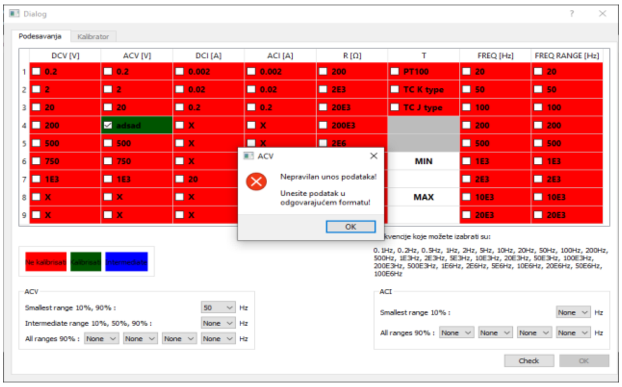
Slika7. Vrednosti i opsezi koji će biti kalibrisani

Kada su izabrane veličine, opsezi i frekvencije na kojima će se vršiti kalibracija digitalnog multimetra, korisnik mora proveriti ispravnost unesenih podataka. To se radi klikom na „Check" taster.

U slučaju da je došlo do nepravilnosti kao što su:

- Nepravilan unos podataka (podtak za opseg veličine se unosi u decimalnom ili u naučnom (eksponencijalnom) formatu 1E3)
- Neizabrane frekvencije ACV i ACI
- Izabrano dva ili više „*intermediate*" opsega

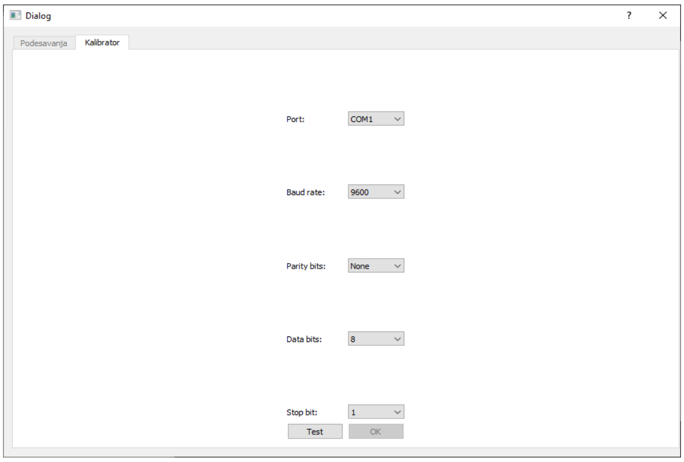
Iskočiće prozor sa upozorenjem o grešci kao na slici 8. Tek kada svi podaci budu prošli proveru „Check" tasterom, moguće je preći na sledeći korak.



Slika8. Greška pri unosu podataka

U poruci greške se nalaze informacije o koloni u tabeli u kojoj je nastala greška i o tipu greške radi lakšeg otklanjanja nastalih grešaka.

Ukoliko provera unesenih podataka prođe bez grešaka, taster „OK“ se odledi. Klikom na taster „OK“ prelazi se na sledeći tab „Kalibrator“ prikazan na slici 9 .



Slika9. Tab za podešavanje parametara komunikacije

U ovom tabu se se postavljaju parametri za uspostavljanje komunikacije kao što su port, brzina prenosa, bit parnosti, bitovi podataka i stop bit. Takođe, u ovom tabu se nalazi taster „Test“ za proveru fizičke povezanosti aplikacije i kalibratora.

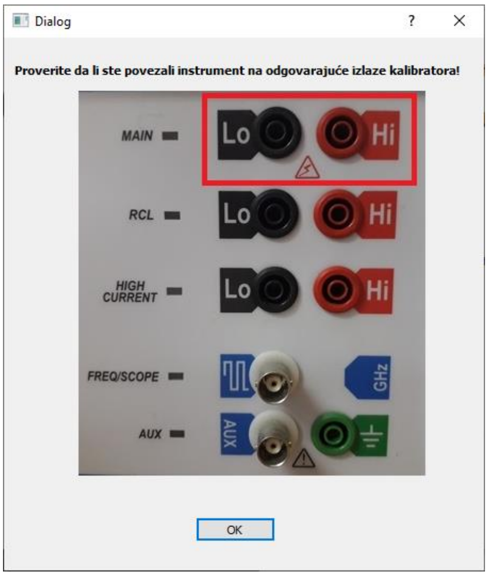
U slučaju da dođe do greške prilikom testa, „Test“ taster će promenuti boju u crvenu i pojaviće se prozor sa porukom o grešci. U tom slučaju korisnik mora proveriti kabl sa kojim su povezani računar i kalibrator ili povezati se na drugi port.

Početak kalibracije

Kada korisnik pokrene komunikaciju, proces kalibracije počinje. Klikom na taster „Podesi“, koje se nalazi na glavnom prozoru, se postavljaju opsezi i vrednosti merene veličine na kalibratoru.

Pre nego što se uključi izlaz kalibratora, pojaviće se prozor sa slikom na kojoj su obeleženi ulazi na kalibratoru na koje se digitalni multimetar treba spojiti.

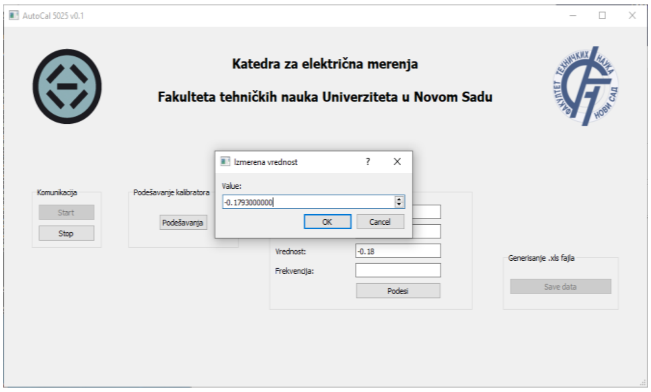
Ovo dodatno osigurava digitalni multimetar od fizičkog oštećenja usled neadekvatnog povezivanja.



Slika10. Povezivanje instrumenta na izlaze kalibratora

Pored ove poruke, u slučaju kada se mere naponi veći od 200 V i struje od 20 A, kalibrator će dati upozorenje zvučnim signalom koji laborantu nagoveštava da obrati pažnju pošto se radi sa visokim naponima i strujama.

Kada korisnik izmeri sve merne veličine i opsege koje je izabrao, otvoriće se prozor sa porukom o tome da je završeno merenje.



Slika11. Prozor za unos izmerenih vrednosti

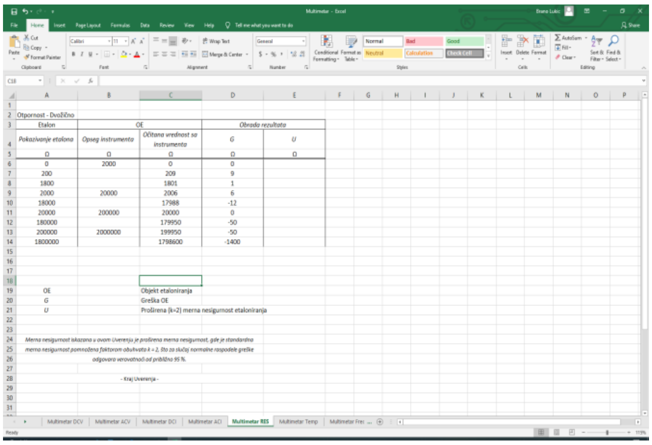
Rezultati merenja

Kada korisnik završi merenja, taster „Save data“ se odledi i klikom na njega se generiše Excel dokument sa izmerenim podacima. Dokument se nalazi u folderu gde je instaliran program. Laborant na kraju merenja unosi dokument sa rezultatima merenja u bazu podataka.

Ime datoteke u kome je sačuvan Excel fajl je datum i vreme kalibracije multimetra. Trenutna verzija aplikacije nema opciju podešavanja putanja za čuvanje podataka.

Excel tabela se sastoji od pet kolona

- Pokazivanje instrumenta
- Opseg instrumenta
- Očitana vrednost sa instrumenta
- G (Greška)
- U (Proširena (k=2) merna nesigurnost etaloniranja koju korisnik ručno unosi)



Slika12. Excel tabela sa rezultatima merenja otpornosti

III. ZAKLJUČAK

U ovom radu je predstavljeno jedno rešenje aplikacije za automatizaciju kalibracije digitalnih multimetara.

Aplikacija doprinosi efikasnosti procesa kalibracije u smislu uštede vremena kao i jednostavnijem podešavanju parametara kalibratora.

Aplikacija je pisana u programskom jeziku Pajton uz korištenje biblioteka za „GUI“ programiranje. Korištenje frejmvorka „QtDesigner“ za izradu „GUI“ aplikacije je dovelo do značajnog smanjenja pisanja koda kao i uštede vremena za izdradu aplikacije.

ZAHVALNICA

Zahvaljujem se profesoru Platonu Sovilju, Nemanji Gazivodi i Đorđu Novakoviću na podršci i pomoći prilikom izrade ovog rada, kao i svim kolegama sa Katedre za električna merenja.

LITERATURA

[1] https://doc.qt.io/qt-5/qtdesigner-manual.html
[2] https://www.timeelectronics.com/multifunction-calibrators/5025e-multifunction-calibrator/
[3] https://www.timeelectronics.com/assets/upload/download_centre/documents/5025-series-brochure_hr.
[4] https://www.euramet.org/Media/docs/Publications/calguides/EURAMET_cg-15__v_2.0_Guidelines_Calibration_Digital_Multimeters.pdf
[5] https://docs.python.org/3.8/

ABSTRACT

Abstract – The paper describes the detailed procedure for the implementation of one of the solutions for automating the process of calibration of digital multimeter using the Times Electronics 5025 calibrator. The main goal of this application is to simplify the procedure of calibration and to save time, besides that it is affordable if we compare it with applications on the market. For the implementation of application the Python programming language with QT framework was used.

Automation of digital multimeter calibration

Branislav Lukić, Đorđe Novaković, Nemanja Gazivoda, Platon Sovilj

Razvoj merno-informacionog sistema za podršku pri etaloniranju temperaturnih sondi

Aleksandar Dimitrijević, Platon Sovilj, *Member, IEEE*, Đorđe Novaković, *Member, IEEE*, Nemanja Gazivoda, *Member, IEEE*

***Apstrakt*— U ovom radu je opisan postupak automatizacije procesa etaloniranja temperaturnih sondi. Prikazan je način izrade aplikacije, implementacija određenjih rešenja za sprečavanje grešaka, kao i algoritam po kojem se kod izvršava. Akcenat je stavljen na pojednostavljenju rada operatera. Uzete su u obzir konsultacije sa više korisnika čime se težilo optimalnom rešenju kako izgleda, tako i funkcionalnosti same aplikacije. Za realizaciju aplikacije korišćen je programski jezik Pajton sa QT frejmvorkom.**

***Ključne reči*—proces etaloniranja; kalibracija; merenje temperature; automatizacija; temperaturne sonde.**

I. Uvod

Primena savremenih uređaja u industriji, medicini, domaćinstvu, poljoprivredi i gotovo svim drugim oblastima zahteva precizno merenje kao i regulaciju temperature. Merenjem temperature dobijamo uvid i vršimo uticaj na rad uređaja, pa samim tim i na funkcionisanje sistema. Merenje temperature omogućava predviđanje havarije i pre nego što se ona desi. Zbog toga je jako bitno da uređaji budu permanentno održavani (kalibrisani, etalonirani), odnosno održavani tako da uvek daju relevantne rezultate.

Sam proces etaloniranja temperaturnih sondi zasniva se na očitavanju dobijenih vrednosti nakon uspostavljenog ravnotežnog stanja zadate temperature i temperaturne sonde. Imajući u vidu da je temperatura sporo promenljiva veličina u sam proces etaloniranja uvodi se automatizacija. Automatizacija olakšava rad korisniku na taj način što ne zahteva njegovo prisustvo tokom odvijanja procesa etaloniranja. Samim tim smanjuje verovatnoću za grešku i ubrzava proces.

II. Termoparovi

Senzore temperature delimo na: ekspanzione, termootporničke, termistore, termoparove, termootporničke senzore od pt-žice, diodne itd. U radu se obrađuju termoparovi i termootpornički senzorori od pt-žice koji imaju najmasovniju primenu.

Aleksandar Dimitrijević – Fakultet Tehničkih Nauka, Novi Sad, Srbija (e-mail: adimitrijevic22@gmail.com).

Platon Sovilj - Fakultet Tehničkih Nauka, Novi Sad, Srbija (e-mail: nemanjagazivoda@uns.ac.rs).

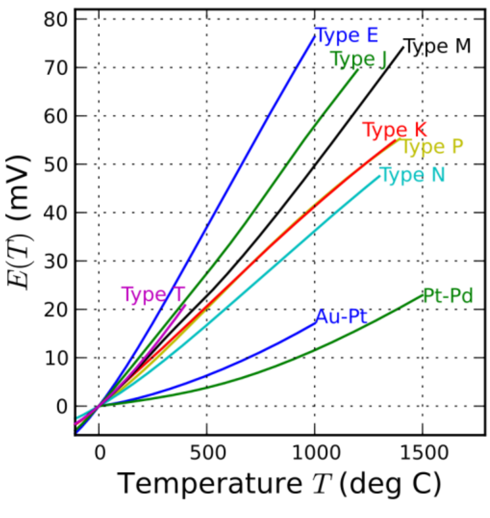
Đorđe Novaković – Fakultet Tehničkih Nauka, Novi Sad, Srbija (e-mail: djordjenovakovic@uns.ac.rs).

Nemanja Gazivoda – Fakultet Tehničkih Nauka, Novi Sad, Srbija (e-mail: platon@uns.ac.rs).

Termoparovi imaju veoma široku primenu u nauci i industriji. Koriste se u sušarama, silosima i u svim segmetima gde je neophodno precizno meriti temperature. Princip rada termoparova zasniva se na *Sibekovom efektu* koji opisuje nastajanje elektromotorne sile između dva različita metala ili njihovih legura. Dobijena elektromotorna sila je proporcijalna razlici između temperature krajeva. Za potrebe očitavanja temperature ne treba nikakav spoljni oblik pobude. Opseg u kojima mogu da mere temperature i izgled karakteristike strogo zavisi od materijala.

Termopar tipa K je najčešći termoelement u upotrebi sa osetljivošću od 28 µV/°C do 41 µV/°C. Jedna od mnogobrojnih prednosti je to što se može naću u širokom spektru opsega temperature. Funkcija kojom se opisuje zavisnost elektromotorne sile u odnosu na temperaturu zadata je od strane NIST-a (*National Instittute of Standards and Tehcnology*). Funkcija je aproksimirana polinomom devetog reda na punom opsegu temperature termopara tipa K (od -200°C do 1100 °C).

$$T = d_0 + d_1 v + d_2 v^2 + ... + d_7 v^7 + d_8 v^8 + d_9 v^9 \quad (1)$$



Sl. 1. Grafički prikaz zavisnosti napona od temperaure termoparova

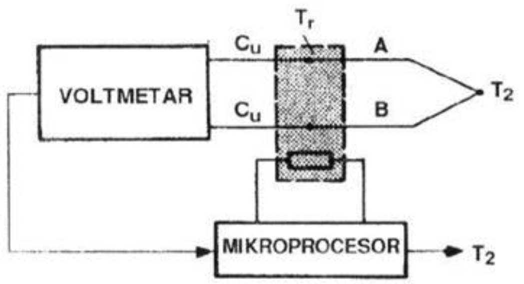
Međutim primenom NIST-ovog polinoma sa deset koeficijenata dobija se promenljiva greška u zavisnosti od izabranog opsega. Ovako dobijena greška pravilnim odabirom koeficijenata u najboljem slučaju može se svesti na maksimalnu vrednost od 0,05 °C. S obzirom da se korišćenjem ovog polinoma ne dobija ista greška na svim opsezima temperature, pribegava se sužavanju opsega na potrebne relacije. Čime se greška znatno smanjuje.

Za proračun vrednosti elektromotorne sile u temperaturu biće korišćena racionalna aproksimacija polinomalne funkcije svedena na odnos dva polinoma manjeg reda umesto jednog polinoma većeg reda.

$$T = T_0 + \frac{(v-v_0)(p_1+(v-v_0)(p_2+(v-v_0)(p_3+p_4(v-v_0))))}{1+(v-v_0)(q_1+(v-v_0)(q_2+q_3(v-v_0)))} \quad (2)$$

Ubacivanjem koeficijenata u jednačinu vezanih za određeni temperaturni opseg greška se svodi na maskimalno moguće odstupanje od ±0,002 °C, što govori o tome da ova metoda daje višestruko tačnije rezultate merenja za specifičan temperaturni opseg.

Imajući u vidu prethodno navedene osobine termopara, za merenje koje će biti realizovano projektom neophodno je izvršiti kompenzaciju hladnog kraja termopara. Naziv metode je softverska kompenzacija hladnog kraja termopara. Preciznije rečeno u laboratorijskim uslovima moguće je obezbediti propisane uslove za realizaciju merenja, međutim sve više se pribegava merenju temperature hladnog spoja (referentnog kraja) sa drugim senzorom temperature (najčešće NTC ili PTC otpornikom). Nakon izmerene temperature hladnog spoja koristi se inverzna transformacija funkcije za proračun temperature na osnovu koje će se dobiti vrednost elektromotorne sile koju hladan spoj generiše. Kompenzacija se vrši samo ukoliko je temperature hladnog kraja različita od 0 °C.



Sl. 2. Šema za kompenzaciju hladnog kraja termopara

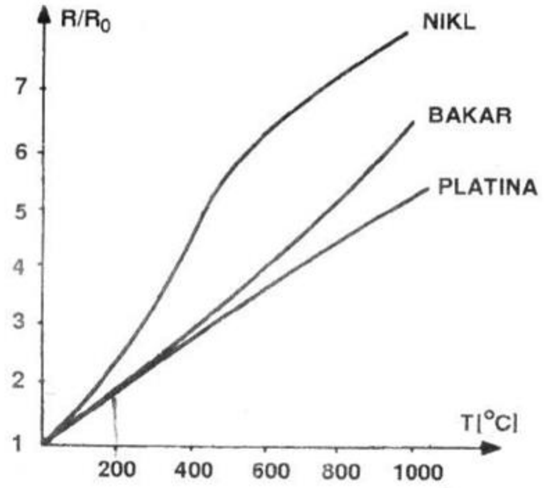
Formula za proračun napona hladnog kraja termopara:

$$v_{cj} = v_0 + \frac{(v-v_0)(p_1+(T_{cj}-T_0)(p_2+(T_{cj}-T_0)(p_3+p_4(T_{cj}-T_0))))}{1+(T_{cj}-T_0)(q_1+q_2((T_{cj}-T_0)))} \quad (3)$$

gde je v_{CJ} proračunati hladni kraj, T_{CJ} temperature hladnog spoja, dok su T_0, v_0, p_1, p_2, p_3, p_4, q_1, q_2 koeficijenti koji zavise od temperaturnog područija merenja.

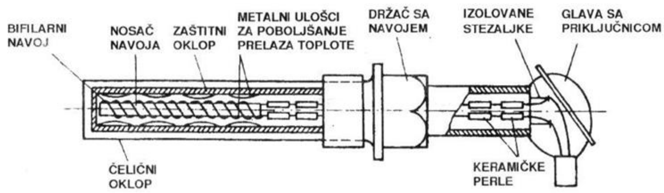
III. Termootpornički senzori

Termootpornički senzori ili RTD senzori su otporni elementi koji menjaju otpor u zavisnosti od temperature. Zbog dobre karakteristike koriste se za precizno merenje temperature sa tačnošću ispod 0,1 °C. Senzor predstavlja obmotanu žicu oko keramičkog ili staklenog jezgra. Žica se izrađuje od platine čija je čistoća 99,999%, ali može biti i od nikla ili bakra.



Sl. 3. Izgled karakteristika termootporničkih senzora u zavisnosti od materijala

Vrednost otpora na temperaturi od 0 °C je 100 Ω ili 1000 Ω pa odatle nazivi Pt100, Pt 1000, Cu 100, Ni 100 itd. Opseg temperature u kojem se senzor može naći je od -200 °C do 850 °C (maskimalno 1500 °C), dok je opseg otpornosti od 10 Ω do 25 kΩ.



Sl. 4. Izgled termootporničkog senzora

Relacija između otpora platinskog senzora i temperature opisuje se *Callendar-Van Dusen* jednačinom (CDV).

Jednačina za proračun otpornosti u korelaciji sa temperaturom manjom od 0 °C:

$$R_{RTD} = R_0(1 + AT + BT^2 + (CT^3(T - 100))) \quad (4)$$

Jednačina za proračun otpornosti u korelaciji sa temperaturom većom od 0 °C:

$$R_{RTD} = R_0(1 + AT + BT^2) \quad (5)$$

Za potrebe projekta neophodna je funkcija koja zavisi od izmerenog otpora. Za dobijanje te funkcije radi se inverzna transformacija jedne od jednačina.

IV. TEHNIČKO REŠENJE

A. *Instrumentacija*

Multimetar FLUKE 8846a je 6 ½ cifarski digitalni multimetar sa dvostrukim ekranom dizajniranim za potrebe laboratorije, rad na terenu i ubacivanje u sistemske aplikacije. Poseduje nekoliko interfejsa za komunikaciju kao što su *RS-232, IEEE 488* i *Ethernet* protokol. *Ethernet* protokol ove uređaje čini idealnim kandidatima za precizna merenja i upotrebu u automatizovanim sistemima.

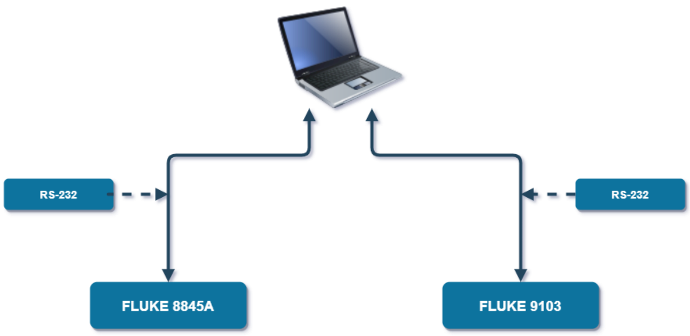
Karakteristike multimetra potrebne za projekat:

- Četvorožično merenje otpornosti
- Merenje jednosmernog napona
- RS-232 komunikacija

Suvi temperaturni kalibrator FLUKE 9103 može se koristiti kao prenosivi instrument za kalibraciju termoparova i RTD temperaturnih sondi. FLUKE 9103 je dovoljno male veličine za upotrebu na terenu takođe i dovoljno tačan za upotrebu u laboratoriji.

Karakteristike kalibratora potrebne za projekat:

- RS-232 protokol
- Univerzalni AC ulaz
- Kontrola brzine skeniranja temperature
- Mogućnost očitavanja u °C ili °F
- Pamćenje osam poslednjih zadataih vrednosti temperature

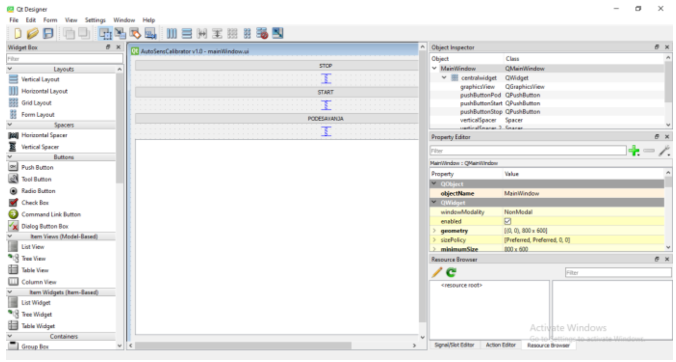


Sl. 5. Blok šema sistema

B. *Pajton biblioteke*

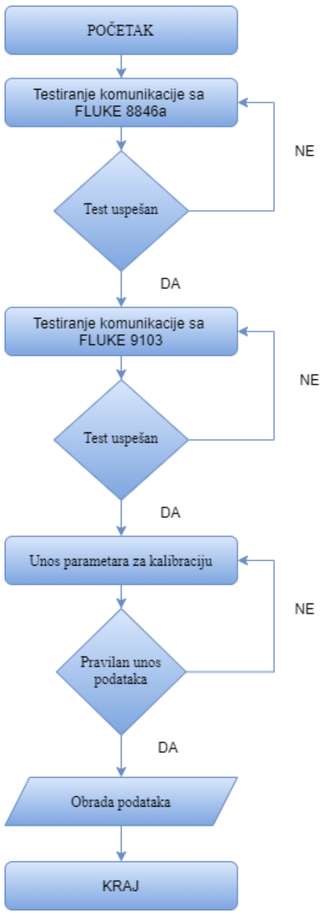
Za razvoj ove aplikacije korišćen je Pajton programski jezik koji je danas veoma popularan. Kroz Pajton se može izvršavati kod koji obavlja jednostavne operacije kao što je sabiranje brojeva pa sve do vrlo kompleksnih operacija koje opisuju ceo jedan sistem.

Pajton biblioteka PyQt5 omogućava korišćenje Qt GUI frejmvorka. Qt je napisan u C++ programskom jeziku. Kada se koristi u simbiozi sa Pajtonom daje znatnu uštedu vremena tokom izrade korisničkog dela aplikacije, a pritom ne gubi na brzini koju pruža C++.



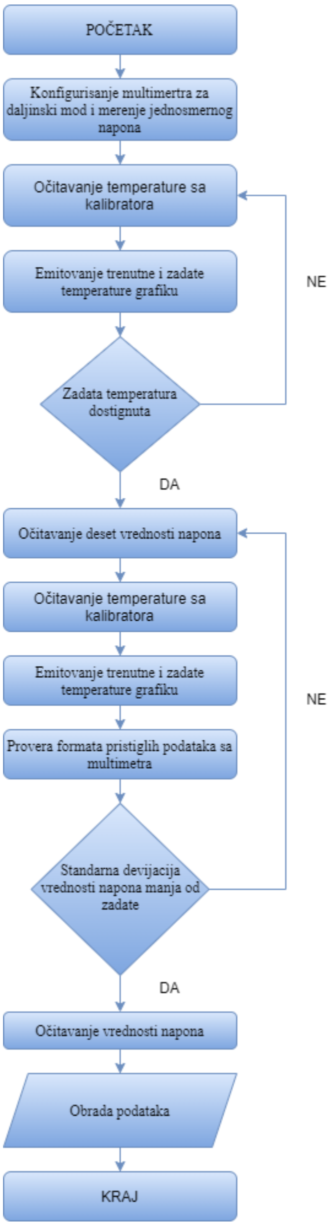
Sl. 6. Izgled glavnog prozora Qt dizajnera

C. *Algoritam aplikacije*

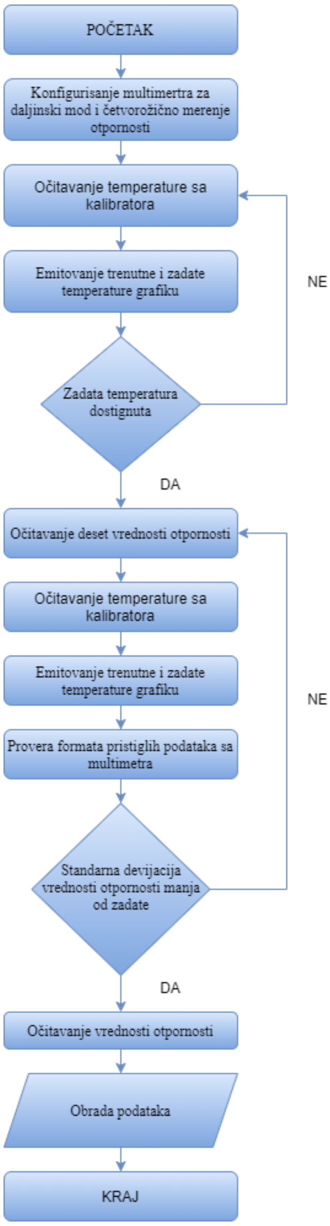


Sl. 7. Algoritam za podešavanje komunikacije i parametara kalibracije

Prikazani algoritam je uvodni algoritam za pravilnu konfiguraciju komunikacije i parametara kalibracije. Algoritam vodi korisnika kroz proces ne dozvoljavajući mu da preskoči neki korak i time na bilo koji način ugrozi merenje. Algoritam se izvršava pritiskom na taster „Podešavanja“. Sl. 10.



Sl. 8. Algoritam toka kalibracije za termopar tipa K



Sl. 9. Algoritam toka kalibracije za temperaturne sonde Pt100 i Pt 1000

Prvi korak algoritma toka kalibracije za termopar tipa K prebacuje multimetar u daljinski režim rada i konfiguriše ga za merenje jednosmernog napona sa automatskom promenom opsega i potrebnom rezolucijom. Tokom kalibracije aplikacija neprekidno dobija informaciju o trenutnoj temperature suvog temperaturnog kalibratora čiju vrednost emituje na grafiku.

Prvi uslov koji vodi ka tačnom očitavanju vrednosti napona koju termopar daje je stabilizacija temperature kalibratora. To se obezbeđuje na taj način što se u određenom vremenskom periodu traži informacija o njegovoj temperaturi. Ukoliko su tri uzastopne vrednosti očitane temperature jednake zadatoj algoritam se nastavlja, u suprotnom proces se ponovo izvršava.

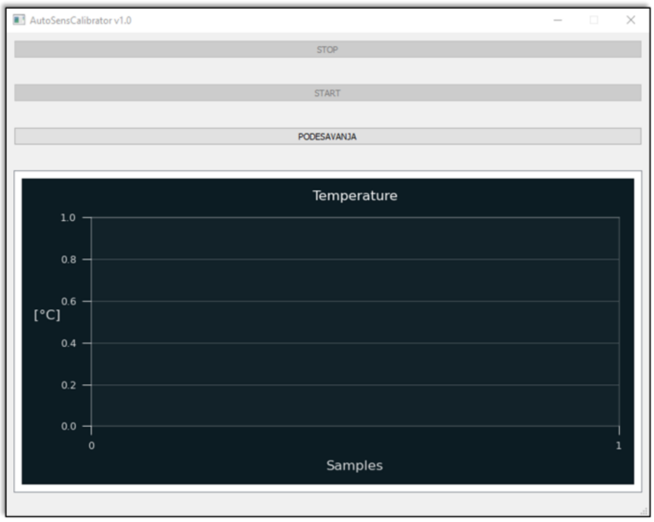
Drugi uslov je provera standardne devijacije deset uzastopnih merenja vrednosti napona dobijenih u kratkom vremenskom periodu za zadato odstupanje.

Prvi korak algoritma toka kalibracije za RTD temperaturne sonde prebacuje multimetar u daljinski režim rada i konfiguriše ga za četvorožično merenje otpornosti sa automatskom promenom opsega i potrebnom rezolucijom. Tokom kalibracije aplikacija neprekidno šalje upit za trenutnu vrednost temperature suvog temperaturnog kupatila koju emituje na grafik. Prvi uslov koji vodi ka tačnom očitavanju vrednosti otpornosti koju RTD daje je stabilizacija temperature kalibratora. To se obezbeđuje na taj način što se u određenom vremenskom periodu traži informacija o njegovoj temperaturi. Ukoliko su tri uzastopne vrednosti očitane temperature jednake zadatoj algoritam se nastavlja, u suprotnom proces se ponovo izvršava.

Drugi uslov je provera standardne devijacije deset uzastopnih merenja otpornosti dobijenih u kratkom vremenskom periodu za zadato odstupanje.

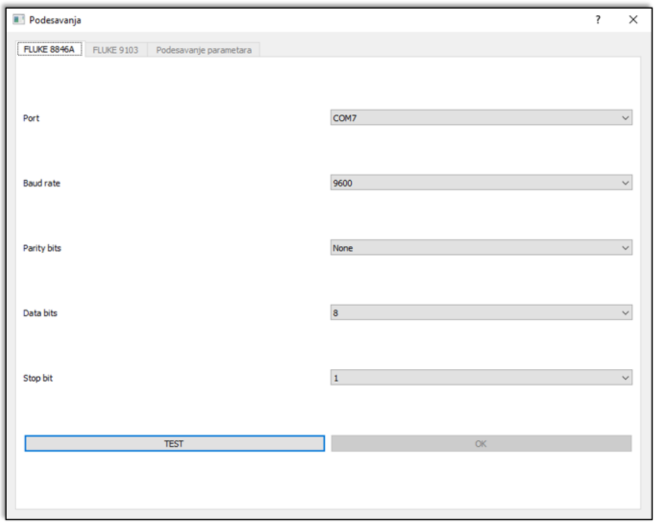
D. Izgled aplikacije

Korišćenjem metoda i podataka iz prethodnih poglavlja izrađena je aplikacija za kalibraciju temperaturnih senzora (AutoSensCalibrator).



Sl. 10. Početni prozor aplikacije

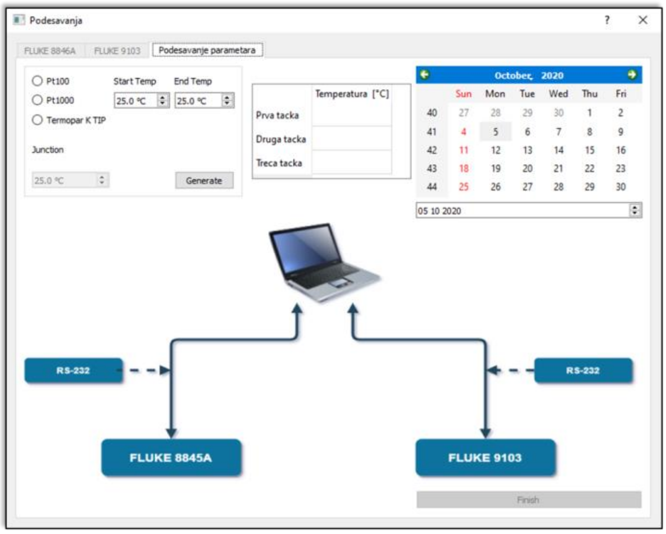
Na Sl. 10. prikazan je početni prozor i elementi koji se na njemu nalaze. Prozor se sastoji od tri tastera: „Start“; „Stop“ i „Podešavanja“. Takođe na istom prozoru je postavljen i grafik (grafik prikazuje trenutnu i zadatu temperaturu kao i kompletnu karakteristiku kalibracije).



Sl. 11. Podešavanje komunikacije sa uređajima

Sl. 11. prikazuje prvi prozor za podešavanje komunikacije sa uređajima. Na tabu FLUKE 8846a prikazana je mogućnost podešavanja parametara za komunikaciju (PORT, BAUD RATE, PARITY BITS, DATA BITS i STOP BIT) i taster „TEST“ za testiranje komunikacije. Pritisak na taster „OK“ se omogućava tek nakon uspešnog završetka testa.

Tab FLUKE 9103 prikazuje podešavanja komunikacije sa suvim temperaturnim kalibratorom FLUKE 9103. Prikazana je mogućnost podešavanja parametara za komunikaciju (PORT, BAUD RATE, PARITY BITS, DATA BITS i STOP BIT).



Sl. 12. Podešavanje parametara za kalibraciju

Na Sl. 13. prikazan je izbor tipa temperaturne sonde, temperaturnog opsega na kojoj će se vršiti kalibracija i vrednost temperature hladnog kraja. Nakon unosa, klikom na taster „Generate“ tabela se popunjava vrednostima tri zadate temperature. Klikom na „Finish“ taster završava se proces podešavanja i aplikacija je spremna za kalibraciju.

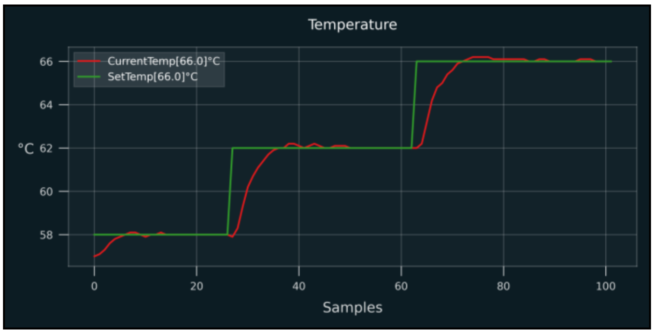
E. Primer izlazne datoteke

Etalon	OB		Obrada rezultata	
Ts	TOEems	TOEconv	G	U
°C	mV	°C	°C	°C
58	1.059	52.64	-5.36	
62	1.202	56.11	-5.89	
66	1.357	59.86	-6.14	
OE	Objekt etaloniranja			
Ts	Temperatura etalonskog sistema			
TOEems	Pokazivanje OE u vrednosti elektromotorne sile			
TOEconv	Konverzijiom dobijeno pokazivanje OE u °C			
G	Greška OE			
U	Proširena (k=2) merna nesigurnost etaloniranja			

Merna nesigurnost iskazana u ovom Uverenju je proširena merna nesigurnost, gde je standardna merna nesigurnost pomnožena faktorom obuhvata k = 2, što za slučaj normalne raspodele greške odgovara verovatnoći od približno 95 %.

- Kraj Uverenja -

Sl. 12. Primer izlazne excel datoteke za termopar tipa K



Sl. 13. Primer izlazne datoteke grafičkog prikaza toka snimanja karakteristike

Nakon završetka kalibracije aplikacija generiše datoteku koja sadrži excel tabelu sa karakteristikama temperaturne sonde i grafikom celokupnog toka kalibrisanja.

V. ZAKLJUČAK

Izrada teme zahtevala je sukcesivan i kontinuiran rad. Sva poglavlja su uzročno posledično povezana. Bez komplesnog i korektnog završetka poglavlja nije se moglo preći na sledeće. To ukazuje na činjenicu da je svako poglavlje na neki način celina za sebe i da zahteva ozbiljnu pažnju i obradu.

Projektovana aplikacija nalazi primenu u svakoj oblasti gde se primenjuju temperaturne sonde ovakvog tipa olakšavajući i davajući preglednost urađene kalibracije. Aplikaciju karakteriše lako rukovanje i kompatibilnost sa velikim brojem platformi.

ZAHVALNICA

Zahvaljujem se profesoru Platonu Sovilju, Đorđu Novakoviću i Nemanji Gazivodi na podršci i pomoći prilikom izrade ovog rada, kao i svim kolegama sa Katedre za električna merenja.

LITERATURA

[1] https://sh.wikipedia.org/wiki/Termopar
[2] http://rdccontrol.com/thermocouples/rtds-101/temperature-calculations/
[3] http://www.mosaic-industries.com/embedded-systems/microcontroller-projects/temperature-measurement/thermocouple/calibration-table#computing-cold-junction-voltages
[4] https://people.vts.su.ac.rs/~gogolak/files/SzenzAkt/SR5SEN.pdf
[5] https://srdata.nist.gov/its90/main/
[6] http://download.flukecal.com/pub/literature/884xa___umeng0300.pdf
[7] https://dam-assets.fluke.com/s3fs-public/2748044_6200_ENG_B_W_1.PDF?1kZ6Cz1sOIhOLrBJgeT2Qn8Oot4YX2gM
[8] FLUKE 8845A/8846A Digital Multimetar-Programers Manual
[9] FLUKE 9103 Dry-Well User's Guide

ABSTRACT

This paper describes the automation of the temperature probes calibration process. Here are presented the method of application development, the implementation of certain solutions for error prevention, and the code execution algorithm. The paper focuses on simplifying the work of operators. In consultation with several users, the goal was to find the optimal solution for the appearance and the functionality of the application itself. Python programming language with QT framework was used for the realization of the application.

Uređaj za ispitivanje tačke rose

Zdravko Gotovac, Rade Peranović, Dragan Pejić, *Member, IEEE*, Platon Sovilj, *Member, IEEE*

***Apstrakt*—U ovom radu upoznaćemo se sa dizajniranjem i programiranjem jednog kompaktnog akvizicionog sistema koji se koristi u svrhe nadgledanja tačke rose. Upotrebom dostupnih i cenovno pristupačnih tehnologija moguće je napraviti akvizicioni sistem koji je značajno jeftiniji od trenutno ponuđenih na tržištu.**

***Ključne reči*—Arduino UNO, LabVIEW, senzorski sistem, Trough hole tehnologija**

I. Uvod

Senzori za merenje vlažnosti vazduha nalaze veliku primenu tamo gde god uticaj vlage iz vazduha čini jednu od karika u lancu merenih veličina. Ovi senzori se naročito koriste u meteorološkim stanicama, laboratorijama, u velikom broju industrijskih grana, u poljoprivredi, u domaćinstvima itd. Upotreba ovih senzora u meteorološkim stanicama ima za svrhu da ljudima da informaciju koliko je zasićenje vazduha vodenom parom kako bi se dalje mogla određivati vremenska prognoza. U labaratorijama senzori vlažnosti vazduha nalaze primenu zato što uređaji koji se tamo nalaze mogu davati pogrešne informacije ukoliko je vlažnost vazduha izvan dozvoljenih granica, ali i hemijski procesi se ne mogu odvijati željenim tokom ukoliko vlažnost vazduha nije zadovoljavajuća. Prehrambena, metalurška, i mnoge druge grane industrije koriste senzore za vlažnost vazduha kako bi svoje proizvodne procese obavljale prema zadatim kriterijumima.

Mjerenje relativne vlažnosti vazduha u ovom radu je realizovano uređajem koji radi na principu određivanja tačke rose. Detaljan način na koji senzor ovog uređaja radi detaljnije je objašnjen u poglavljima II i III.

Ali pre nego što se počne sa opisom o diskretnim delovima čitavog sistema treba reći kako čitav sistem funkcioniše.

Sistem meri temperaturu vazduha i temperaturu temperaturno reglulisanog peltijeovog elementa, pri čemu se „zamagljenost“ reflektivnog elementa drži na 70%. Zamagljenost je prethodno određena u procesu autokalibracije maksimalnim zagrijavanjem, odnosno hlađenjem peltijeovog elementa, pri čemu su maksimumi funkcija očitavanja optičkog senzora, kao i samo očitavanje „zamagljenosti“ od 70% reflektivnog elementa, odnosno ogledala.

Projekat je osmišljen tako da se njegova realizacija obavlja u više diskretnih delova, odnosno podeljen je na nekoliko delova na kojima se može relativno nezavisno raditi.

Zdravko Gotovac – Fakultet Tehničkih Nauka, Novi Sad, Srbija (e-mail: zdravko.gotovac@uns.ac.rs).

Rade Peranović – Fakultet tehničkih nauka, Novi Sad, Srbija (e-mail: rade.peranovic123@gmail.com).

Dragan Pejić – Fakultet tehničkih nauka, Novi Sad, Srbija (e-mail: pejicdra@uns.ac.rs).

Platon Sovilj – Fakultet tehničkih nauka, Novi Sad, Srbija (e-mail: platon@uns.ac.rs)

Delovi od kojih se projekat sastoji su:

A. dizajniranje i izrada temperaturnog akvizicionog kola

B. dizajniranje i izrada optičkog akvizicionog kola

C. dizajniranje i izrada napajanja peltijeovog elementa

D. pisanje programa za akviziciju podataka

Projekat je podjeljen u diskretne delove da bi se jednostavnije mogle vršiti izmene prilikom izrade, kao i zbog potencijalne mogućnosti unapređivanja pojedinačnih delova rada, kao i rada u celini.

II. Izrada projekta

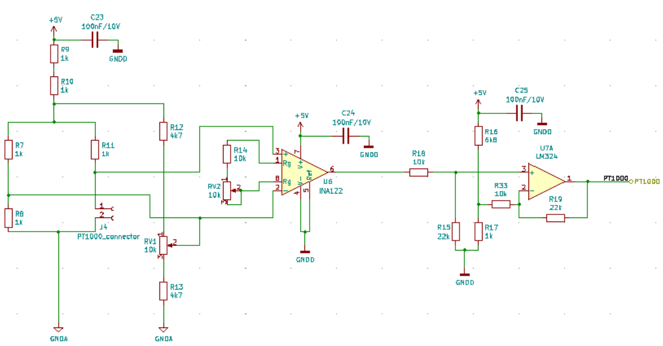
A. Dizajniranje i izrada temperaturnog akvizicionog kola

Jedan od najbitnijih delova čitovog sistema jesu hardverska kola za akviziciju temperature. Osnovna ideja za kola koja se koriste za merenje temperature jeste da se akvizicija izvrši sa više različitih senzorskih modula.

Za merenje temperature peltijeovog elementa koristi se kolo za akviziciju koje u sebi sadrži platinski otpornik PT 1000 otpornosti 1000 Ω.

Njegova otpornost je temperaturno zavisna, odnosno sa temperaturom, raste i njegova otpornost. Naravno i inverzna relacija važi.

PT 1000 se nalazi povezan u Vitstonov most sa otpornicima nazivne otpornosti 1 kΩ, dok se na red sa mostom, na napajanju nalaze povezana dva otpornika od 1 kΩ, kojima se ograničava struja koja prolazi kroz PT 1000.



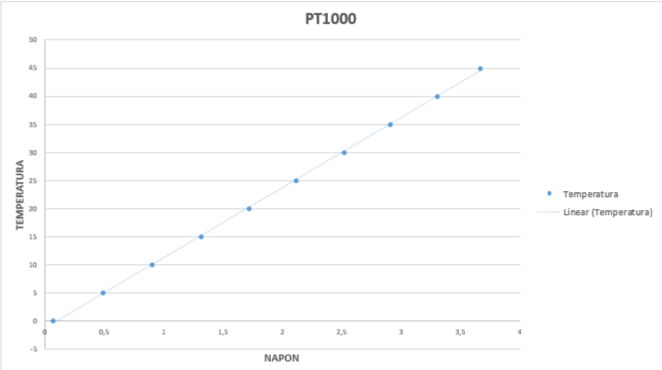
Sl. 1. Šema dela kola za akviziciju temperature

Pored toga, u paraleli sa vitstonovim mostom, nalazi se povezano kolo sačinjeno od tri redno vezana otpornika, od kojih je jedan potenciometar čiji je kraj na kom se nalazi klizač povezan na tačku Vitstnonovog mosta na kojoj se nalaze otpornici nazivne otpornosti 1 kΩ.

Njegovom upotrebom uspostavljamo ravnotežu mernog mosta kada je temperatura PT 1000 0 °C. To treba da se radi jer tačne vrednosti otpornika odstupaju od njihovih nazivnih i do 5%, što bi u ovom slučaju predstavljalo 50 Ω, a uzimajući u obzir da je se temperatura PT 1000 menja ± 3,9 Ω za svaki °C [1].

Naponske signale koje dobijamo iz Vitnstonovog mosta povezujemo na INA122 [2] instrumentacioni pojačavač, a potom i na LM324 [3] koji se koristi u konfiguraciji instrumentacionog pojačavača.

Ovako akviziciono kolo daje kao rezultat signal čija se vrednost kreće od 73 mV do 3,663 V. Odnos vrednosti temperature i izlaznog signala može se videti na grafiku zavisnosti.

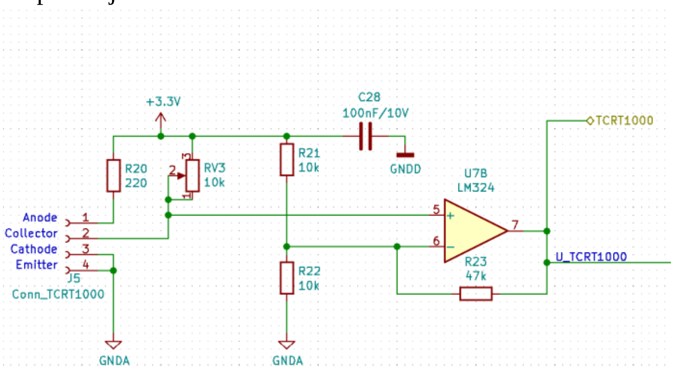


Sl. 2. Signal koji se dobija na izlazu iz dela za akviziciju temperature

B. *Dizajniranje i izrada optičkog akvizicionog kola*

Kao što je bilo neophodno dizajnirati kolo za akviziciju temperature, bilo je neophodno dizajnirati i napraviti optički detektor koji će moći da prima i obrađuje informacije o zamagljenosti ogledala.

Da bi se moglo i početi sa tim bilo je neophodno odlučiti se za djelove kola kojima se vrši akvizicija optičkih signala, a zbog relativno malog kućišta izabran je TCRT1000 [4]. Reč je o uparenoj IR LED i foto tranzistoru.



Sl. 3. Šema dela za akviziciju optičkog signala

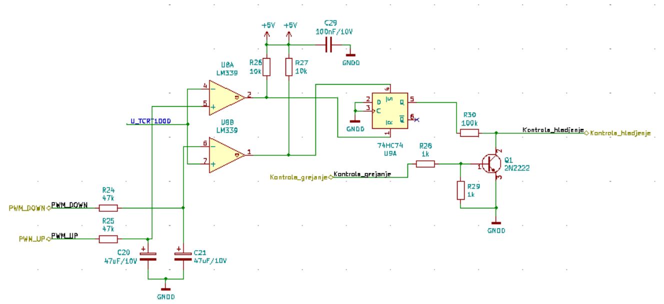
Opto element je povezan na napajanje od 3,3 V i na svom kolektoru daje stujni signal koji se kasnije vodi na LM324 i u konfiguraciji neinvertujućeg operacionog pojačavača pojačava i vodi na komparator.

Tako se signal pretvara u povorku pravougaonih signala koja se dalje vodi na set/reset flip-flop koji se u kombinaciji sa softverskim modulom koristi za upravljanje H-mostom.

C. *Dizajniranje i izrada napajanja Peltijeovog elementa*

Za objašnjenje napajanja peltijeovog elementa, potrebno je prvo objasniti napajanje H-mosta.

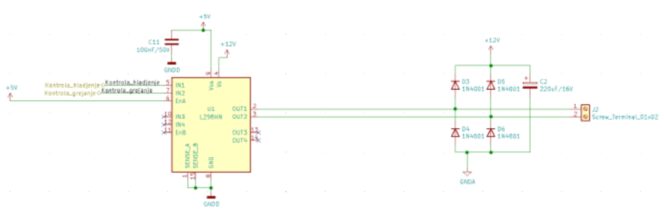
Da bi se omogućilo njegovo ispravno funkcionisanje, odnosno da bi se obezbjedile i mogućnost rashlađivanja gornje površine pletijeovog elemetna, a samim tim i grijanja reflektivnog elementa, ogledala, potrebno je upravljati H-mostom tako da se menja polaritet na njegovim krajevima za napajanje.



Sl. 4. Šema dela korištenog za hardversku kontrolu H-mosta

Time upravljamo dovođenjem signala na dva pina H-mosta predviđena za potrebnu kontrolu, pri čemu se na jedan od pinova dovodi signal iz opto elementa, dok se na drugi dovodi digitalni signal iz mikrokontrolera kojim se vrši akvizicija signala, kao i samo upravljanje čitavim sistemom.

Sam H-most se napaja sa 5 V, koji su obezbeđeni sa napojne jedinice koja je prethodno napravljena.

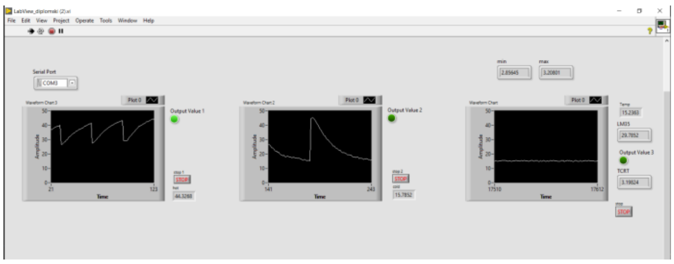


Sl. 5. Šema povezivanja povezivanja H-mosta

D. *Pisanje programa za akviziciju podataka*

Za akviziciju i kontrolu čitavog sistema koristi se Arduino UNO [5] razvojna platforma. Da bi se ona uspešno mogla koristiti potrebno je isprogramirati čitav sistem.

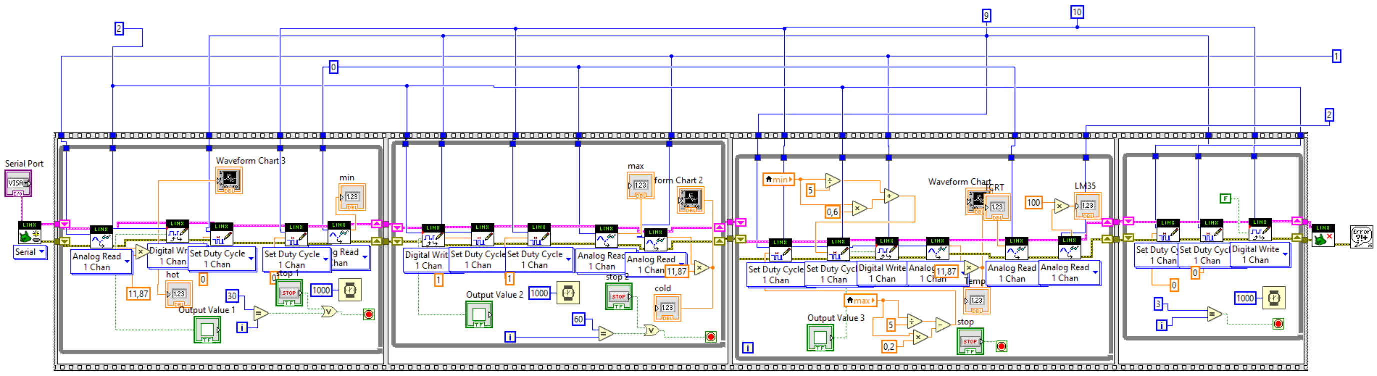
Softversko okruženje koje se koristi za pisanje firmvera potrebnog za funkcionisanje čitavog sistema je LabVIEW [6] softverski paket, koji je besplatno dostupan studentima.



Sl. 6. Izgled programa za akviziciju i obradu podataka

Za komunikaciju između računara na kom je pisan firmver, korištena je LINX [7] besplatna softverska nadogradnja.

Firmver je napisan tako da se program kaskadno izvršava, pri čemu se prvo izvršava deo koda predviđen za autokalibraciju akvizicionog kola, nakon čega se deo za ispitivanje tačke rose izvršava sve do trenutka kada se obustavlja rad uređaja. Na kraju se izvršava deo koda koji je napisan za potrebe resetovanja čitavog akvizicionog sistema, kao i komunikacije mikrokontrolera sa računarom.



Sl. 7. Izgled programa za akviziciju i obradu podataka

III. ZAKLJUČAK

Na osnovu izloženog može se zaključiti da je izrada jednog sistema za akviziciju tačke rose raltivno jednostavan postupak, ukoliko su dostupna potrebna znanja i alati za izradu istog.

Naravno kada se uzme u obzir uloženi trud za čitav projekat, finansije potrebne za izradu postaju počinju da rastu, ali kada se izrađeni uređaj uporedi sa široko dostupnim uređajima koji se nalaze na tržištu, njegova konačna cena ostaje i dalje pristupačnija.

ZAHVALNICA

Zdravko Gotovac i Rade Peranović žele da se zahvale Draganu Pejiću, Đorđu Novakoviću, Platonu Sovilju i Zoranu Mitroviću na korisnim diskusijama i savetima, kao i na obezbeđivanju potrebne opreme prilikom izrade rada.

LITERATURA

[1] Inst Tools. https://instrumentationtools.com/rtd-standards/ [accessed Feb. 22, 2021]

[2] INA122 AMPLIFIER Datasheet. https://datasheetspdf.com/pdf/255474/Burr-BrownCorporation/INA122/1 [accessed Feb. 22, 2021]

[3] LMx24-N, LM2902-N Low-Power, Quad-Operational Amplifiers. https://www.ti.com/lit/ds/snosc16d/snosc16d.pdf [accessed Feb. 22, 2021]

[4] TCRT1000, TCRT1010 Reflective Optical Sensor with Transistor Output. https://www.vishay.com/docs/83752/tcrt1000.pdf [accessed Feb. 27, 2021]

[5] Getting Started with Arduino UNO | Arduino. https://www.arduino.cc/en/Guide/ArduinoUno [accessed Feb. 29, 2021]

[6] What is LabVIEW? – NI. https://www.ni.com/en-rs/shop/labview.html [accessed Mar. 11, 2021]

[7] LINX [LabVIEW MakerHub]. https://www.labviewmakerhub.com/doku.php?id=libraries:linx:start [accessed Mar. 11, 2021]

ABSTRACT

In this work we shall make ourselves familiar with hardware and software development of compact acquisition system used for purposes of finding dew point. Main idea on which this work is based on is usage of existing, free or relatively inexpensive hardware and software solutions to create significantly less expensive system than ones available at open market.

Dew point acquisition device

Zdravko Gotovac, Rade Peranović, Dragan Pejić, Platon Sovilj

Predlog implementacije komunikacionih i kontrolnih metoda u konceptu Industrije 4.0

Zdravko Gotovac, Marjan Urekar, *Member, IEEE*

***Apstrakt*—U ovom radu upoznaćemo se sa dizajniranjem i implementacijom komunikacionih i kontrolnih metoda korišćenjem šrioko rasprostranjenih tehnologija. Čitava ideja predstavlja logičan iskorak u uključivanje Industrije 4.0 u proizvodni process, kao i indirektno unaprijeđivanje uređaja i radnog okruženja koji su realni dio proizvodnog procesa.**

***Ključne reči*—Industrija 4.0, IoT, senzorski sistem, baze podataka, proces proizvodnje**

I. Uvod

Da bi se bolje shvatio rad u cjelini, prvo je neophodno, ili poželjno, shvatiti zbog čega se implementira u proces proizvodnje različite funkcionalnosti koje pružaju protokoli definisani unutar Industrije 4.0 [1]?

Odgovor na jedno ovakvo pitanje je dosta kompleksan, i da bi mogli da bi se mogao pružiti zadovoljavajući odgovor na njega neophodno je kritički sagledati potenciajlne dobre, i loše osobine koje pruža Industrija 4.0 i IoT [2]. Pored toga potrebno je i idejno definisati one djelove procesa proizvodnje koje imaju potencijal da budu unaprijeđeni.

Kako se sve više insistira na kontroli i nadgledanju radnog prostora, jedan od očiglednih načina za unaprijeđenje tih radnih uslova je nadgledanje temperature, vlažnosti, vazdušnog pritiska unutar radnog prostora. Najjednostavniji način implementacije bio bi upotreba pametnih telefona koji na sebi sadrže mnoštvo senzora, kao i čipove koji omogućavaju međusobnu komunikacju, i vezu sa internetom.

Pored toga osposobljavanje djelova unutar procesa proizvodnje za međusobnu komunikaciju upotrebom uređaja koji se ne koriste direktno unutar procesa proizvodnje predstavljalo bi izvestan napredak za one uređaje koji ne posjeduju moderne protokole za komunikaciju.

II. Problemi Industrije 4.0 prilikom implementacije u proces proizvodnje

Prije nego sto bi i počeli bilo kakvu raspravu o implementaciji Industrije 4.0, trebalo bi skrenuti pažnju na probleme, i potencijalne grerške koje se javljaju tokom njene implementacije. Čitava priča o umrežavanju sistema proizvodnje zvuči sjajno, ali kao i bilo šta drugo, ona nije bez mana.

Prilikom prikupljanja ogromnih količina podataka može da dođe do zasićenja sistema, što može predstavljati ogroman problem. Postavlja se pitanje da li su svi primljeni i obrađeni podaci potrebni, odnosno čemu bi oni služili? Kako razdvojiti podatke krucijalne za funkcionisanje sistema od „digitalne buke“ koju bi predstavljali beznačajni podaci?

Naravno postoji i velika opasnost ugrožavanja lične privatnosti i slobode preuzimanjem i obradom podataka. Da li bi zaista postojala potreba da šef radnika unutar fabrike barata informacijama o njihovom dnevnom unosu kalorija, ili možda krvnom pritisku? Čemu služi informacija o istoriji pretraživanja interneta jednog radnika? Za takve podatke sigurno postoje lica i kompanije koje bi platile da dođu do njih, da bih ih koristila za marketinške ili neke još manje etičke svrhe.

Naravno kako je pitanje bezbjednosti ličnih podataka trenutno aktuelna tema na globalnom nivou [3], potrebno je postaviti određena ograničenja na implementaciju idejnog riješenja. Zadatak zaštite ličnih podataka, uprkos tome što je značajan, nije jedini na koji treba obratiti pažnju, isto tako potrebno je smanjiti proizvodnju i upotrebu novih uređaja koji bi se koristili samo u svrhe komunikacije, jer bi u slučaju njihovog zastarijevanja predstavljali elektronski otpad.

Ideja bi bila upotreba aplikacije, napisane za više mobilnih platformi, koja bi korisniku pružala mogućnost odobravanja i autentikacije prilikom odabira podataka koje želi da djeli i unosi u sistem. Slični protokoli, naravno opštije prirode već bivaju implementirani u novim verzijama opše dostupnih mobilnih platformi.

Implementacije Industrije 4.0 unutar procesa proizvodnje trebalo bi ograničiti samo na djelove koji bi bili bitni unutar samog procesa proizvodnje.

III. Implementacija Industrije 4.0

A. *Indirekatno umrežavanje sistema proizvodnje*

Jedna od najbitnijih osobina koju ljudska bića posjeduju, i osobina koja je omogućila napredak čovječanstva je mogućnost komunikacije. I danas čovjek, po prirodi komunikativno biće pokušava da unaprijed načine razmjene podataka, informacija i ideja. Počevši od razmjene kamenih ploča na kojima su urezani pa sve virtuelnih online konferencija način razmene podataka postaje sve sofisticiraniji.

Takav način razmišljanja trebalo bi primjeniti i unutar prozvodnog procesa koristeći se protokolima Industrije 4.0. zašto ne omogućiti mašinama da komuniciraju međusobno?

Zagušenje proizvodnog procesa tokom proizvodnje su česte pojave koje dovode do redova čekanja i zbog asinhronizacije čitavog procesa proizvodnje stvaraju dodatne probleme i dodatna zagušenja proizvodnje.

Zdravko Gotovac – Fakultet Tehničkih Nauka, Novi Sad, Srbija (e-mail: zdravko.gotovac@uns.ac.rs).

Marjan Urekar – Fakultet Tehničkih Nauka, Novi Sad, Srbija (e-mail: urekarm@uns.ac.rs).

Šta ako bi diskretni dijelovi unutar sistema posedovali komunikacione linije koje bi omogućavale automatskoj kontroli da slobodno modulišu brzinu proizvodnje unutar sistema. Naravno trebalo bi postaviti kao nominalnu brzinu onu koja bi odgovarala kapacitetima same proizvodne linije, pa u zavisnosti od ispunjenosti tih kapaciteta davati umanjenu brzinu u odnosu na onu nominalnu. Tako bi se otklonile i greške nastale ljudskim faktorom, odnosno potencijalnim nedostatkom krucijalnog osoblja koje se nalazi na određenim delovima proizvodne linije.

Pored kontrole same proizvodnje u sistem bi trebalo uvesti i monitoring na nivou mašine. Šta ako jedna od mašina na djelu proizvodne linije nije u nominalnom režimu rada? Kako to može uticati na ostatak proizvodnje ununtar proizvodne linije? Pretpostavimo da se čitav sistem proizvodnje nalazi na sistemu za napajanje koji je jedinstven za fabriku, i pretpostavimo da jedna od mašina nije u nominalnom stanju rada. To bi značilo za mašinske strojeve da u slučaju povećanja struje u odnosu na nominalnu, dolazilo bi do pregrijevanja sistema napajanja. Ukoliko bi sistem napajanja posjedovao ograničenja, bilo da su to ograničenja u vidu snage koju on može da da, ograničenja u vidu struje ili termalna ograničenja, ona bi svejedno uticala na krajnje funkcionisanje sistema. Zbog jedne mašine koja nije u nominalnom režimu čitav sistem bi mogao da se zaustavi, ili još gore, ošteti. To bi mogli da spriječimo očitavanjem parametara te mašine, njene temperature, struje koju preuzima iz sistema, napona na njenim priključcima, kao i potencijalnim djelovima električnih kola koja se u njoj nalaze, a koja ne bi funkcionisala na ispravan način.

Informacije prikupljene na ovaj način davale bi određeni dio slike o stanju sistema proizvodnje, ali ta bi slika bila potpunija ukoliko bi se uvela dodatna linija za prikupljanje informacija, čiji bi potencijalni izvor bili radnici koji se nalaze unutar sistema proizvodnje.

Danas rijetko ko ne posjeduje pametni mobilni telefon, a uzimajući u obzir nivo napretka te tehnologije, trebalo bi je upotrebiti na nabolje moguće načine.

Jedan pametni telefon sadrži nekoliko senzorskih sistema, čija bi očitavanja mogla da posluže razumjevanju trenutnog stanja sistema, kao i predviđanju njegovog budućeg funkcionisanja. Senzori temperature, mikrofoni koji se mogu koristiti kao senzori zvučnog intenziteta i frekvencijskog odziva, barometri i sl [4]. Iskorišteni na pravi način produžili bi potencijalni radni vijek mašina koje se koriste u sistemu proizvodnje, a isto tako bi unaprijedili i radno okruženje za radnike.

Šta ako mašina radi unutar prostorija koje nisu adekvatno klimatizovane, da li će to dovesti do prestanka njenog rada? Ukoliko vazdušni pritisak nije odgovarajućeg nivoa, da li će to usporiti radnike unutar procesa proizvodnje? Da li bi se mikrofon mobilnog telefona mogao iskoristiti za praćenje zvukova koji su indikatori potencijalnih problema unutar sistema proizvodnje? Na ova i sliča pitanja potrebno je odgovoriti implementacijom odgovarajućeg sistema za nadgledanje procesa proizvodnje.

Zamislimo situaciju, unutar proizvodnog pogona posjedujemo mašinu relativno velikih dimenzija, tako da je neophodno da na njoj radi više radnika. U slučaju da svako od njih posjeduje pametni telefon, mogao bi da instalira aplikaciju koja bi mjerila temperaturu i odzive određenih frekvencijskih spektara, i na osnovu tih informacija pružala informacije o stanju mašine, kako radnicima, lokalno na njihovim pametnim uređajima, tako i glavom računaru slanjem podatka na glavni server. Ako bi se desio neki kvar zbog kog bi mašina ispustila zvuk koji nije ununtar spektra koji može da čuje ljudsko uho, aplikacija bi obavjestila radnika o potencijalnoj opasnosti koja se tu javlja. Ako bi se mašina na jednom od svojih djelova zagrijavala dovoljno sporo radnik možda to ne bi mogao da primjeti, ali aplikacija bi bilježila podatke tokom vremena pa bi postojali grafici na kojima bi se jasno moglo vidjeti da jedan od djelova sistema ne funkcioniše na ispravan način.

Pored toga, na osnovu mrežnih tačaka preko kojih se povezuje aplikacija, mogla bi se pružiti informacija o trenutnoj lokaciji radnika unutar fabrike, što bi značajno popravilo bezbjednosni status tog radnika u slučaju da dođe do kvara unutar sistema proizvodnje koji je opasan po njegov život. Na sličan način bi se mogao primjeniti i algoritam koji će pokazati da li radnik ispunjava zadatke za koje je zadužen.

B. Modularni sistem proizvodnje

Naravno modernizaciju proizvodnog sistema koristeći se definisanim protokolima i pravilima Industrije 4.0 trebalo bi iskoristiti na različite načine, i jedna od ideja je izmjena samog sistema proizvodnje.

Podaci koji se prikupljaju ne moraju biti iskorišteni samo za organizaciju radnika, nego i mašina kojima oni upravljaju. Pretpostavimo da se u fabrici proizvode automobili, i na jednom djelu proizvodne trake se proizvode motori, odnosno platforme koje se ugrađuju u automobile, u drugom djelu šasije automobila, trećem vrata, četvrtom bi se obavljalo farbanje određenih djelova automobila, petom ugradnja neohpodne elektronike i enterijera, i u finalnom bi se postavljali točkovi na automobile.

Da li je isti nivo prioriteta za radnom snagom u svim procesima? Naravno da nije, točkove je na kraju najlakše postaviti, i to zahtjeva najmanje vremena, ali da li je isti slučaj sa proizvodnjom motora? Šta ako se javi greška u procesu proizvodnje motora, odnosno platforme? Ne bi bilo smisla obustavljati ostale djelove procesa ukoliko se u njima ne nalazi greška, bar ne dok se ne otkloni greška u procesu proizvodnje platforme. Radnici bi i dalje mogli da nastave sa proizvodnjom šasija i njihovim farbanjem, dok bi se ugradnja elektronike i kompletnog enterijera automobila obustavila zbog potencijalnih problema. Znači li to da se može indefinitivno nastaviti sa proizvodnjom šasija za automobile? Naravno da ne, uvjek postoje određena logistička ograničenja, od kojih bi jedno od najbitnijih bilo skladištenje finaliziranih šasija.

Algoritmi za obradu podataka, odnosno alogritmi za mašinsko učenje bi se mogli uključiti u sistem proizvodnje [5], gde bi uzimajući u obzir prirodu kvarova koji bi se nalazili unutar diskretnih sistema proizvodnje preuzimali ulogu organizacije i raspoređivanja zadataka u preostalim delovima sistema.

C. Izbacivanje faktora nesigurnosti unutar procesa proizvodnje

Jedan od glavnih faktora koji unose nesigurnost u čitav sistem prizvodnje su ljudi koji se nalaze unutar tog sistema, počevši od radnika koji obavljaju osnovne poslove unutar čitavog sistema, pa sve do onih koji se nalaze na najvišim upravnim pozicijama. Te faktore je neophodno ograničiti i svesti na minimum, da bi se povećao nivo optimizacije čitavog proizvodnog sistema.

Programi koji bi pratili efikasnost radnika, grupa, kao i onih koji upravljaju grupama bi doveli do smanjenja grešaka koje nastaju kada iste te zadatke obavljaju odgovorni ljudi koji daju subjektivna mišljenja i ocjene. Naravno to ne bi značilo da je potrebno zanemariti te subjektivne ocjene, naprotiv, svi obrađeni podaci bi se koristili da suplementiraju pravljenju radnog kartona koji bi ocjenjivao radnike.

Možda se problem nije javio zbog toga što radnik ne izvršava radne zadatke na zadovoljavajući način, nego zbog toga što su mu dati zadaci koji nisu u skladu sa opisom i prikladnom obukom posla kojim se bavi. Isto tako se može javiti problem prevelikog opterećenja radnika, ili korištenja neadekvatnih alata na koje je ograničen. To upućuje na grešku koja se javila od strane uprave koja nije izvršila pravilu raspodjelu radnih zadatake na dostupne resurse.

Greške takve prirode bi mogle da se smanje obradom razmjenjenih podataka, od strane radnika upravi, i obrnuto, pri čemu bi se pravile jasne kategorije koje bi razvrstavale nivo prioriteta određenim priloženim informacijama.

IV. ZAKLJUČAK

Na osnovu priloženog može da se vidi da je implementacija “pametnog proizvodnog sistema” realno ostvariva mogućnost.

Naravno, prije svega, prilikom implementacije jednog takvog sistema moralo bi da se vodi računa o bezbednosti informacija koje bi mogle da ugroze lica koja imaju vlasništvo nad bitnim privatnim informacijama.

Sem toga potrebno je implementirati akvizicioni sistem koji bi na pogodan način selektovao i čuvao odabrane podatke, jer bi to bio prvi korak ka pravljenju robusnog sistema koji bi omogućio prelaz iz prethodno uspostavljenih industrijskih standarda u one koje traži Industrija 4.0.

I ono što je najbitnije je da posmatranjem čitave problematike iz ugla upotrebe široko dostupnih tehnologija omogućilo bi ostvarivost finansijski i logistički pristupačnog riješenja. Tako da bi i implementacija od strane kompanija doživjela veću rasprostranjenost.

LITERATURA

[1] “Industry 4.0: the fourth industrial revolution – guide to Industrie 4.0. ” i-Scoop https://www.i-scoop.eu/industry-4-0/. [accessed Jun. 12, 2021]

[2] “internet of things” [IoT Agenda] https://internetofthingsagenda.techtarget.com/definition/Internet-of-Things-IoT. [accessed accessed Jun. 12, 2021]

[3] “Case Law Digest 2021: Transfers of personal data to third countries” [EUROPEAN DATA PROTECTION SUPERVISOR] https://edps.europa.eu/data-protection/our-work/our-work-by-type/court-cases_en. [accessed accessed Jun. 15, 2021]

[4] “ARDUINO SCIENCE JOURNAL” [Arduino] https://www.arduino.cc/education/science-journal. [accessed accessed Jun. 16, 2021]

[5] Ndeye Lo, Jean-Marie Flaus, Olivier Adrot. Review of Machine Learning Approaches In Fault Diagnosis applied to IoT System. ICCAD 2019 - 3rd International Conference on Control, Automation and Diagnosis, Jul 2019, Grenoble, France. ff10.1109/ICCAD46983.2019.9037949ff. ffhal-02344344

ABSTRACT

This paper introduces the design and implementation of the communication and control methods, alongside the use of widely available technologies. The idea represents a logical step forward in the incorporation of Industry 4.0 in manufacturing process, as well as indirect improvement of production appliances and work environment, which are part of real world manufacturing process.

Implementation of communication and control methods in concept of Industry 4.0

Zdravko Gotovac, Marjan Urekar

Sistem za merenje i regulaciju temperature u zamrzivačima za čuvanje Pfizer-BioNTech COVID-19 vakcine

Milan Šaš, *Student Member, IEEE*, Bojan Vujčič, *Member, IEEE* i Dragan Pejić, *Member, IEEE*

***Apstrakt*—U ovom radu biće prikazano rešenje za merenje i regulaciju temperature u zamrzivačima za čuvanje Pfizer-BioNTech COVID-19 vakcine na -70 °C. Rad je baziran na Arduino Nano platformi koja se koristi za obradu podataka koje dobija iz kola koje meri napon na PT1000 senzoru temperature. Dalje, te podatke obrađuje i prikazuje na četvorocifrenom sedmosegmentnom displeju. Cilj rada jeste da se pokaže mogućnost projektovanja i implementiranja sistema za merenje i regulaciju temperature koja nije u standardnom opsegu merenja. Opseg temperature koji se ovim sistemom meri je od -100 °C do 0 °C.**

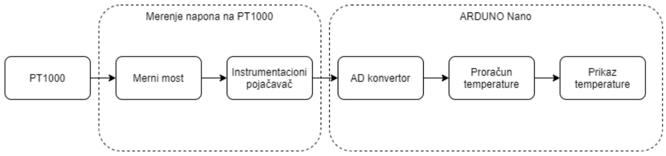
***Ključne reči*—COVID-19, Pfizer-BioNTech, vakcina, merenje temperature, PT1000, INA122, Arduino Nano, merni most.**

I. Uvod

U prethodnom vremenu smo svedoci posledica Covid-19 virusa koji je korenito promenio način života mnogih od nas i stavio zdravstvene sisteme svih zemalja na test. Kao rezultat razvoja tehnologije i medicine, koji prate jedan drugog u stopu, imamo vakcine za razne bolesti koje su iskorenjene u prošlosti. Sa pojavom novih vakcina protiv Covid-19 pojavila se nada da će i ova pošast proći što pre. Kompanije Pfizer (SAD) i BioNTech (Nemačka) [1] su razvile vakcinu protiv Covid-19 sa visokim stepenom efikasnosti ali i jednom manom: potrebno je čuvati vakcinu na -70 °C. Ovo nije standarda temperatura za čuvanje vakcina pa je potrebno projektovati sistem za merenje i regulaciju koji omogućava transport vakcine na toj temperaturi kako bi ona došla u sve delove sveta.

Sistem koji je projektovan i koji će biti prikazan u ovom radu bavi se upravo ovim problemom ali ima za cilj da se ovaj problem reši tako da samo rešenje bude „*low cost*“ rešenje, kako bi moglo da se lako i jeftino primeni u praksi. Na slici 1 data je blok šema celog sistema.

Milan Šaš - Fakultet tehničkih nauka, Univerzitet u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (email: milansas@uns.ac.rs).
Bojan Vujčić – Fakultet tehničkih nauka, Univerzitet u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (email: bojanvuj@uns.ac.rs).
Dragan Pejić – Fakultet tehničkih nauka, Univerzitet u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (email: pejicdra@uns.ac.rs).

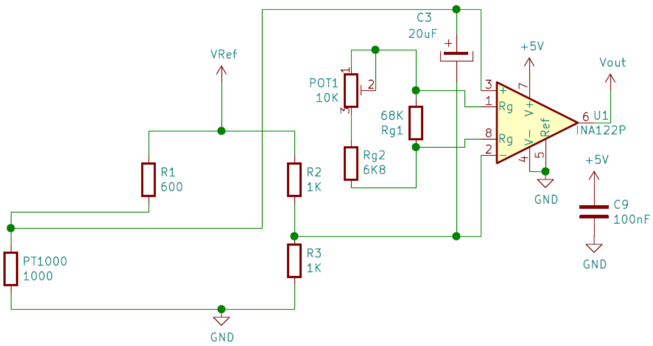


Sl. 1. Blok šema sistema

II. Komponente sistema

A. *Merenje temperature PT1000 sondom*

Temperatura se meri pomoću PT1000 sonde [2], trožično, tako što se terminali sonde dovode u merni most i tada sonda postaje jedan od otpornika u mostu. Ostali otpornici su prilagođeni nameni ovog sistema. Pošto znamo da otpornost PT1000 na 0 °C iznosi 1000 Ω a na -100 °C iznosi 602.6 Ω, te vrednosti su uzete kao vrednosti otpornika u mostu. Most se napaja sa 2.5 V (*Vref*) a napon sa merne dijagonale se vodi na instrumentacioni pojačavač INA122 [3] koji napon iz mosta pretvara u napon u opsegu od 0 V do *Vref*. Na slici 2 je dat prikaz mosta i instrumentacionog pojačavača.

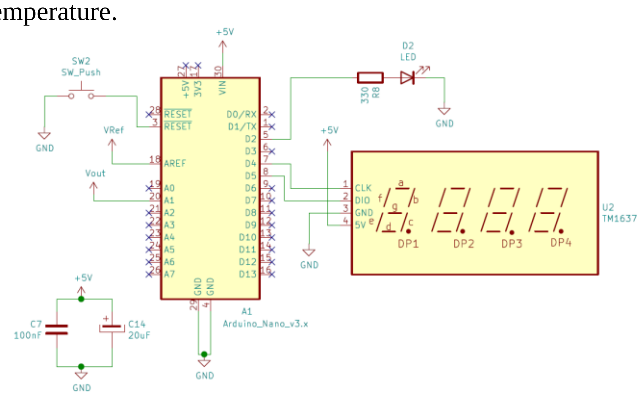


Sl. 2. Merni most i instrumentacioni pojačavač [3]

Instrumentacioni pojačavač se napaja unipolarno sa 5 V a otpornik Rg, koji definiše pojačanje instrumentacionog pojačavača, ima otpornost od 67,2 kΩ. Kako bi dobili što približniju vrednost otpornika Rg koristi se trimer od 10 kΩ i dva otpornika od 68 kΩ i 6,8 kΩ koji su povezani kao što je prikazano na slici 2. Podešavanjem trimera imamo fino zadavanje otpornosti Rg. Dodatno, postavljeni su kondenzatori C3 od 20 uF i C9 od 100 nF koji se koriste za stabilisanje napona na ulazu instrumentacionog pojačavača i napona napajanja, respektivno.

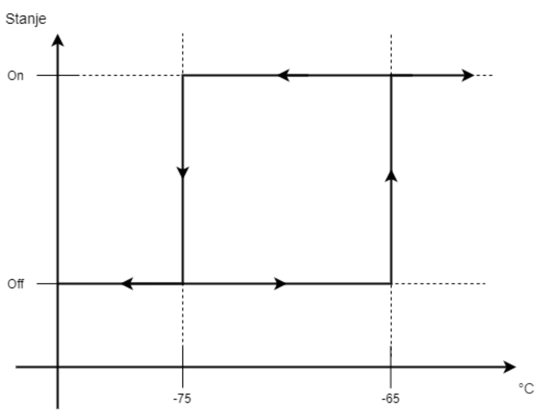
B. Arduino Nano i obrada podataka

Baza sistema je Arduino Nano [4] razvojno okruženje zbog niske cene, malih dimenzija i velike dostupnosti. Na slici 3 dat je izgled i povezivanje Arduino Nano sistema sa sedmosegmentnim displejom TM1673 [5] za prikaz temperature.



Sl. 3. Arduino Nano [4] i displej TM1637 [5]

Na analogni ulaz A1 se dovod izlaz iz instrumentacionog pojačavača. Takođe, koristi se mogućnost zadavanja referentnog napona za AD konvertor i na AREF pin Arduino Nano sistema dovodi se napon *Vref* kojim se, ujedno, napaja merni most. Kako bi imali mogućnost resetovanja sistema povezan je taster između GND i RESET pina Arduina. Kako bi ostvarili regulaciju temperature, pin D2 je postavljen kao izlazni i na njega povezana crvena LED koja daje informaciju o radu kompresora ili nekog drugog uređaja koji može da snizi temperaturu u objektu. Realizovano je histerezisno upravljanje tako što su u kodu koji se nalazi na Arduinu postavljene granice histerezisa na -65 °C i -75 °C. Kada se temperatura popne iznad -65 °C izlaz D2 na Arduinu će preći u visoko stanje. Kada temperatura padne ispod -75 °C izlaz D2 na Arduinu će preći u nisko stanje. Na slici 4 je dat izgled histerezisa prema kome se vrši upravljanje, odnosno kako se menja stanje na pinu D2.

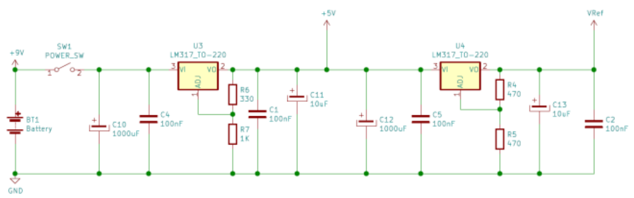


Sl. 4. Histerezis

Na Arduinu je implementiran algoritam koji vrši očitavanje pina A1 tako što izvodi AD konverziju. AD konvertor koji je sastavni deo Arduino platforme je 10-bitni pa ceo opseg merene temperature dobijamo u opsegu od 0 do 1023. Kada se izvrši AD konverzija potrebno je prebaciti dobijenu vrednost AD konverzije u temperaturu. Više o formuli koja vrši to pretvaranje biće rečeno u sledećem odeljku. Primenom formule dobijamo vrednost temperature u °C koja se prikazuje na četvorocifrenom sedmosegmentnom displeju koji je sa dve linije povezan sa Arduinom. Pin CLK je povezan na pin D4 a pin DIO je povezan na D5. Kako bi rad sa ovim displejom bio olakšan koristi se Arduino biblioteka istog imena kao i dispelj.

C. Napajanje

Napajanje uređaja je projektovano tako da možemo koristiti izvod od 9 V DC, najčešće baterija ili drugi tip izvoda DC napona. Na slici 5 je data šema napajanja.



Sl. 5. Napajanje uređaja

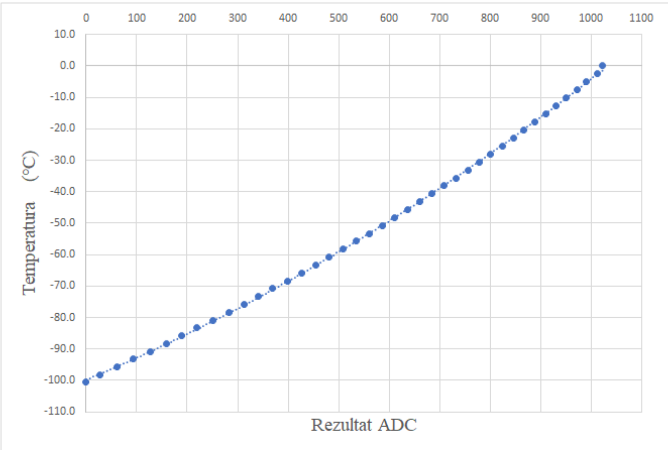
Potrebno je postići dva naponska nivoa, od 5 V i 2.5 V, a to postižemo tako što koristimo dva LM317 [6] naponska regulatora konfigurisana tako da nam daju prethodno pomenute naponske nivoe. Nivo od 5 V se koristi za napajanje instrumetacionog pojačavača, Arduina i displeja a nivo od 2.5 V se koristi kao referenca za AD konvertor i napajanje mernog mosta. Ovde je bitno da se most napaja sa istom referencom koja se dovodi na AD konvertor kako bi dobili raciometrijsko merenje.

III. OBRADA PODATAKA

Iz tabele otpornosti za senzor PT1000 možemo videti da njegova otpornost na 0 °C iznosi 1000 Ω a na -100 °C iznosi 602.6 Ω. Kako što je rečeno u prethodnom odeljku, ove vrednosti su uzete kao vrednosti otpornika u mostu. Sada nam je bitna zavisnost otpornosti od temperature. Postoje dva pristupa rešavanju ovog problema. Prvi pristup jeste da prvo gledamo zavisnost otpora od temperature na senzoru, posle toga da gledamo zavisnost napona u mostu od promene otpora i na kraju da gledamo zavisnost rezultata AD konverzije od promene napona. Pošto je ovo previše komplikovano za naše potrebe mi smo se odlučili za drugačiji pristup ovom problemu i direktno smo posmatrali zavisnost rezultata AD konverzije od temperature. Kolo je sastavljeno i spremno za prvo testiranje. Pošto nismo u mogućnosti da postignemo temperaturu od -100 °C morali smo da simuliramo otpornost koju bi dobili na temperaturama u već pomenutom opsegu. Otpornost smo simulirali dekadnom kutijom MA 2115 proizvođača Metrel [7], prikazanu na slici 6.

Sl. 6. Otporna dekada MA 2115 [7]

U kolo je povezana dekada i sada smo spremni za merenje. Procedura je sledeća: dekadu postavimo na otpornost od 600 Ω i povećavamo je do 1000 Ω u koracima od 10 Ω a zapisujemo rezultat AD konverzije. Rezultat ovog merenja je dat na grafiku 1:

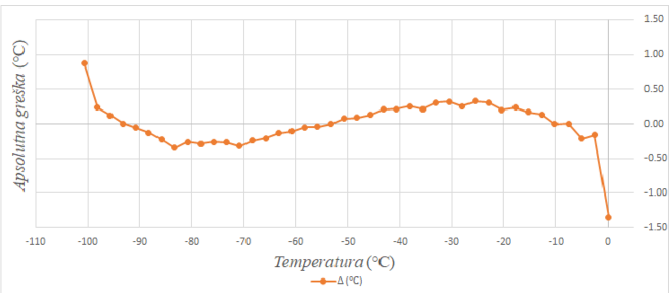


Gr. 1. Zavisnost rezultata AD konverzije od temperature

Aproksimiranjem prikazane zavisnosti kvadratnom jednačinom dobijamo jednačinu (1):

$$Temperatura(°C) = 28,536 * 10^{-6} * ADC^{2} + 0,067041 * ADC - 99,756 \quad (1)$$

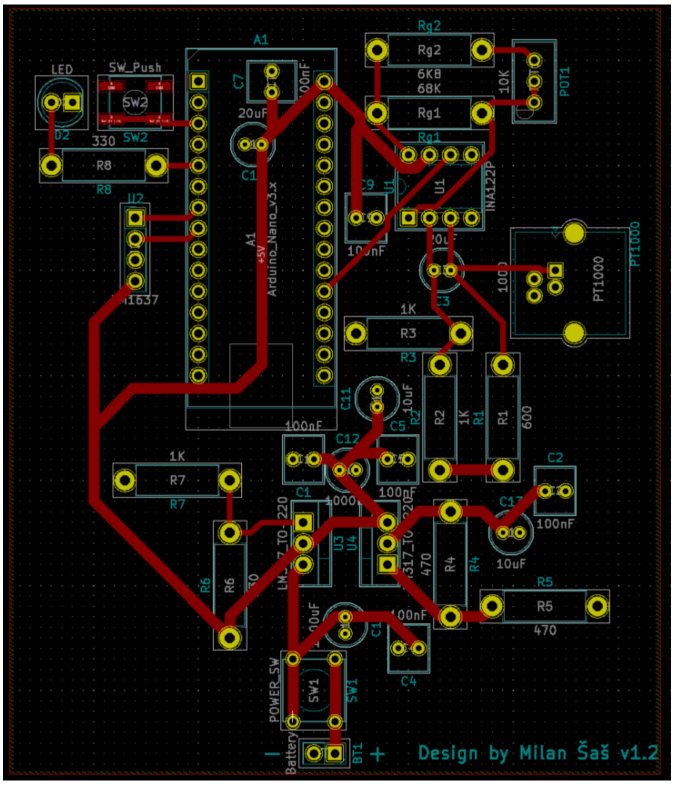
Ovo je jednačina koja je deo algoritma u Arduinu i na osnovu koje dobijemo vrednost temperature. Ukoliko se porede izmerene vrednosti sa zvaničnom tabelom senzora PT1000 možemo da vidimo da apsolutna greška ne prelazi 0.5 °C, ako zanemarimo krajeve opsega. Analiza apsolutne greške je data na grafiku 2.



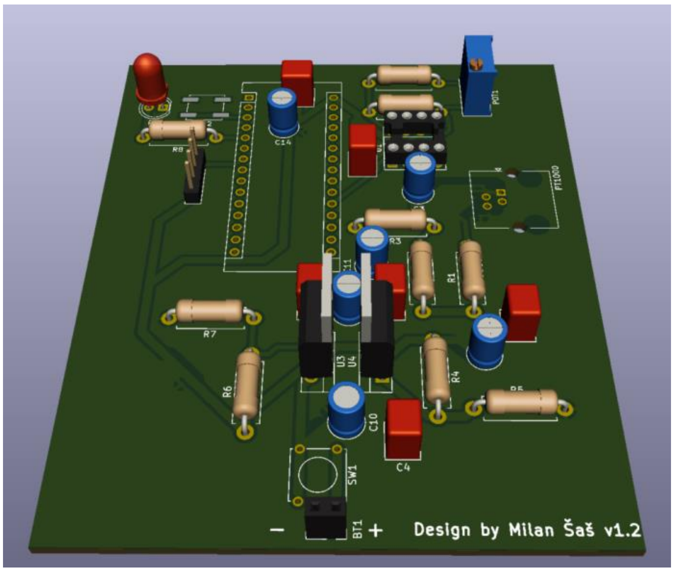
Gr. 2. Apsolutna greška merenja temperature

IV. IZRADA UREĐAJA

Kada smo testirali sistem i utvrdili da zadovoljava naše kriterijume prešli smo na izradu samog uređaja. Projektovanje PCB-a je rađeno u program KiCad [8] a jedno od mogućih rešenja PCB-a je prikazano na slikama 7 i 8:



Sl. 7. PCB izgled



Sl. 8. PCB izgled 3D

Konačni izgled uređaja je dat na slici 9:

Sl. 9. Izgled samog uređaja (levo) i uređaj povezan na dekadnu kutiju (desno)

V. ZAKLJUČAK

Na kraju rada možemo zaključiti da je sistem uspešno realizovan i da zadovoljava prvobitno planirane potrebe. Sam sistem se može primeniti i u drugim poljima rada gde je potrebno meriti i regulisati temperaturu u datom opsegu, kao što je transport zamrznute ribe i morskih plodova sa primorja u kontinentalne krajeve kako bi se ta riba servirala sveža u suši barovima i ostalim restoranima. Zanimljivo je primetiti da pojava samozagrevanja koji se javlja kod PT100 i PT1000 senzora ovde ne predstavlja problem zato što samozagrevanje doprinosi rastu temperature i imamo ranije aktiviranje releja koji dalje uključuje uređaj koji snižava temperaturu pa ovde možemo da kažemo da za samozagrevanje važi rečenica: *"It's not a bug, it's a feature"*. Uređaju je dato ime ŠVP-21 u čast svih autora ovog rada. Dalji planovi podrazumevaju unapređenje postojećeg sistema kako bi se smanjila greška u krajevima opsega kao i greška na čitavom opsegu merenja temperature. Za samu regulaciju greška na krajevim opsega ne pravi problem, ukoliko bi želeli manju grešku možemo uzeti polinom višeg stepena. Takođe, dalji planovi podrazumevaju i testiranje uređaja u nekom od temperaturnih kupatila laboratorije za metrologiju na Fakultetu tehničkih nauka u Novom Sadu.

LITERATURA

[1] Pfizer, "PFIZER-BIONTECH COVID-19 VACCINE," [Online]. Available: https://www.pfizer.com/products/product-detail/pfizer-biontech-covid-19-vaccine.

[2] K. M. AS. [Online]. Available: https://www.kongsberg.com/globalassets/maritime/km-products/product-documents/tsiec751_ce.pdf.

[3] T. Instruments, "Literature," [Online]. Available: https://www.ti.com/lit/ds/symlink/ina122.pdf.

[4] Arduino. [Online]. Available: https://www.arduino.cc/en/uploads/Main/ArduinoNanoManual23.pdf.

[5] D. cafe, "TM1637 datasheet," [Online]. Available: http://www.datasheetcafe.com/tm1637-datasheet-pdf/.

[6] STMicroelectronics. [Online]. Available: https://www.st.com/resource/en/datasheet/lm317.pdf.

[7] Mertel, "MA 2115 and MA 2115 S Decade resistors," [Online]. Available: https://www.metrel.si/en/shop/LABS/ma-2115-and-ma-2115-s-decade-resistors.html.

[8] KiCad, "Home page," [Online]. Available: https://www.kicad.org.

ABSTRACT

This paper will present a solution for measuring and regulating the temperature in Pfizer-BioNTech COVID-19 vaccine freezers at -70 °C. The work is based on the Arduino Nano platform used to process the data obtained from the circuit that measures the voltage on the PT1000 temperature sensor. Further, it processes and displays this data on a four-digit seven-segment display. The aim of this paper is to show the possibility of designing and implementing a system for measuring and regulating temperature that is not in the standard measuring range. The temperature range measured by this system is from -100 °C to 0 °C.

Temperature measurement and control system in Pfizer-BioNTech COVID-19 vaccine freezers

Milan Šaš, Bojan Vujčić, Dragan Pejić

Edukativni pristup enkriptovanom prenosu podataka u embedded i frontend razvojnim okruženjima

Ivan Gutai, Member, IEEE, Prof. dr Platon Sovilj, Member, IEEE, Marina Subotin, Member, IEEE, Marjan Urekar, Member, IEEE, Jelena Milojević, Member, IEEE, Milica Mitrović, Member, IEEE

***Apstrakt*—Pre IIoT-a (The Industrial Internet of Things), embedded programiranje i frontend programiranje nisu mogli da se nađu ni u istoj rečenici. Hardver koji je omogućio neprimetnu integraciju ove dve kompleksne oblasti je Espressif-ov ESP32 MCU (MicroController Unit). Tek kada je broj uređaja povezanih na internet dostigao značajnu cifru, u fokus je došla bezbednost podataka. ESP-NOW je Espressif-ova tehnologija za bežični prenos podataka. Prenos može biti enkriptovan i obezbeđuje bezbednu kumunikaciju između više ESP32. HTTPS (Hypertext Transfer Protocol Secure) je protokol koji povećava bezbednost na internetu. Ovaj rad predstavlja uputstvo za konfigurisanje razvojnog okruženja za ESP32, ESP-NOW primer, primer HTTPS servera i prikaz programerske prakse za upravljanje greškama. Kao dodatak prikazan je i prilagodljivi grafički korisnički interfejs IIoT uređaja. U ovom trenutku postoji mnogo putanja u embedded i frontend programiranju. U ovom radu je izabrana putanja: ESP32 za hardver i C++ za firmver. JavaScript, HTML5 i CSS3 su neizbežan deo modernih industrijskih uređaja, pa je dat primer korišćenja JavaScript Highcharts biblioteke. Korišćena kombinacija hardvera i softvera košta manje od 10$, što čini konfiguraciju pogodnom za zemlje u razvoju. Highcharts biblioteka je vlasnički softver, ali u edukativne svrhe se može koristiti u okviru Creative Commons (CC) Attribution-Non-Commercial licence.**

***Ključne reči*— IIoT; embedded programiranje; frontend programiranje; ESP32; ESP-NOW; HTTPS; prilagodljivi dizajn; SPIFFS; C++; JavaScript; HTML5; CSS3; Highcharts; Web Bazirani Merno-Akvizicioni Sistemi; JSON.**

I. Uvod

Pametni uređaji su postali deo naše svakodnevice, a podjednako ih koristimo i kao alat i kao nešto što se može nazvati hobi projektom. Broj takvih uređaja i njihovih funkcionalnosti se svakodnevno uvećava. U takvim okruženjima moramo obratiti pažnju na bezbednost informacija i ne smemo zaboraviti dobre programerske i inženjerske prakse. Danas svi imamo jednaku mogućnost da razvijemo prototip IIoT uređaja, koji će biti deo naše kućne Wi-Fi mreže, a uređaj možemo kontrolisati preko web pretraživača. Ovaj rad daje niz smernica, sa namerom da čitaocima ubrza ulazak u svet embedded programiranja i/ili web programiranja. Dato je praktično uputstvo kako se kreira jedan IIoT uređaj. Najteži deo na početku je izbor pravog hardvera i odgovarajućeg skupa tehnologija. Autori su izabrali: ESP32[1], C++, JavaScript, HTML5 i CSS3. ESP-NOW [2] tehnologija omogućava povezivanje velikog broja ESP32 uređaja, koji komuniciraju međusobno preko Wi-Fi-ja. Čist tekst je podložan izmenama u web aplikacijama u toku prenošenja preko Wi-Fi-ja. Ukoliko dobijemo priliku da nešto enkriptujemo, to treba odmah da uradimo. ESP-NOW podaci mogu biti enkriptovani preko LMK (Local Master Key), koji mora da se slaže i na prijemnicima i na predajnicima. Dodavanje senzora i releja je zasebna oblast i čitaoci mogu da biraju između raznih open-hardware rešenja i vlasničkih alternativa kao što su proizvodi Mikroelektronike. Nakon završetka sa hardverskim delom, koristi se SPIFFS (Serial Peripheral Interface Flash File System) memorija, za postavljanje (eng. deploy) web aplikacije. SPIFFS sadrži web aplikaciju koja je kreirana sa modernim i besplatnim alatima. Jedina razlika je u hosting-u i aplikacija se ne nalazi na tipičnom web serveru, već se nalazi na ESP32.

Ivan Gutai – Fakultet tehničkih nauka, Novi Sad, Srbija (e-mail: gutai@uns.ac.rs).
Platon Sovilj – Fakultet tehničkih nauka, Novi Sad, Srbija (e-mail: platon@uns.ac.rs).
Marina Subotin – Fakultet tehničkih nauka, Novi Sad, Srbija (e-mail: marina.bulat@uns.ac.rs).
Marjan Urekar – Fakultet tehničkih nauka, Novi Sad, Srbija (e-mail: urekarm@uns.ac.rs).
Jelena Milojević – Fakultet tehničkih nauka, Novi Sad, Srbija (e-mail: jmilojevic@uns.ac.rs).
Milica Mitrović – Fakultet tehničkih nauka, Novi Sad, Srbija (e-mail: m.mitrovic@uns.ac.rs).

II. Konfigurisanje integrisanog razvojnog okruženja

Ceo proces se započinje instalacijom Arduino IDE (Integrated Development Environment) [3].

ESP32 treba dodati na listu postojećih razvojnih sistema: "Arduino IDE, File, Preferences, Additional Boards Manager URLs [4]."

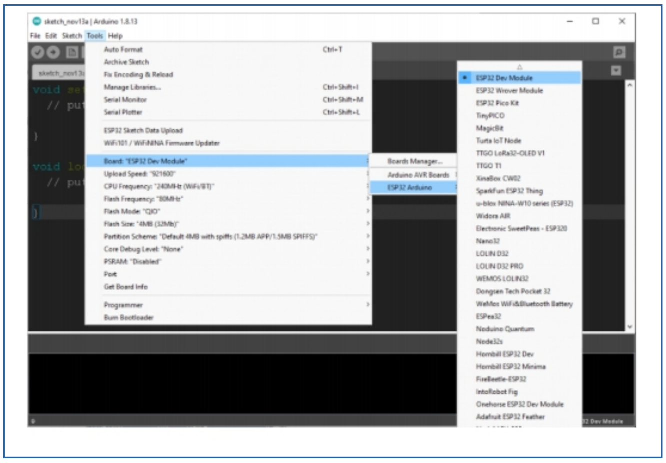
Proširivanje liste postojećih razvojnih sistema sa ESP32: "Tools, Board, Boards Manager, esp32 i Install". Nakon uspešne instalacije, "ESP32 Arduino" će se pojaviti na listi dostupnih razvojnih sistema. Biće dostupno mnoštvo ESP32 razvojnih ESP32 razvojnih sistema, uključujući i "ESP32 Dev Module".

Dodavanje "ESP32 Sketch Data Upload" opcije omogućava postavljanje fajlova na SPIFFS. Možemo to zamisliti kao postavljanje frontend koda na ESP32 sistem. Dodatak (eng. plugin) [5] za postavljanje koda je potrebno kopirati unutar Arduino IDE foldera. Dodatak je potrebno iskopirati u sledeći direktorijum: "C:\Program Files (x86)\Arduino\tools\ESP32FS\tool\esp32fs.jar".

Tamna tema za Arduino IDE je dostupna na [6]. Fajlovi od nove teme treba da zamene postojeće, u folderu: "C:\Program Files (x86)\Arduino\lib\theme". Poželjno je da se fajlovi od originalne teme takođe sačuvaju.

Fajlovi koji su potrebni, da bi uopšte bio moguć rad sa ESP32 su [7]. Navedene fajlove je potrebno otpakovati u lokalni Arduino folder: "C:\Users\Ivan Gutai (odgovarajuće korisničko ime)\Documents\Arduino\hardware\espressif\esp32". Treba napomenuti da je "esp32" folder zapravo preimenovani "arduinoesp32-master" folder.

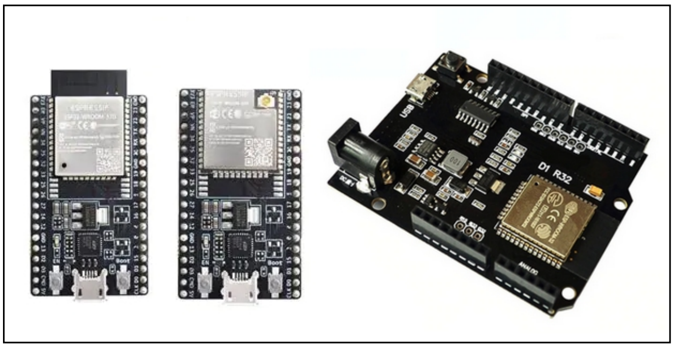
Ukoliko ESP32 razvojni sistem nije vidljiv u Device Manager-u, potrebno je instalirati drajvere za programer. Najpopularniji su CH340 [8] i CP210x [9]. Nakon uspešne konfiguracije, razvojno okruženje izgleda kao na slici 1.



Sl. 1. Arduino integrisano razvojno okruženje, konfigurisano za ESP32.

III. Izbor odgovarajućeg hardvera

ESP32 ima više varijacija. U ovom radu je korišćen ESP32-DevKitC-32D [10] sa integrisanom bakarnom antenom. Da je bilo potrebe da pojačavamo Wi-Fi signal, upotrebili bi ESP32-DevKitC-32U [11], sa zasebnom antenom. Ukoliko je neko navikao da radi sa Arduino Uno, postoji ESP32 koji fizički veoma sličan, a reč je o verziji Wemos D1 R32 [12]. Slika 2 prikazuje sva tri spomenuta tipa ESP32 razvojnih sistema.



Sl. 2. ESP32 razvojni sistemi: ESP32-DevKitC-32D, ESP32-DevKitC- 32U i ESP32 Wemos D1 R32.

IV. Istraživanje ESP-NOW tehnologije

Postoji kompletno uputstvo za konfigurisanje ESP-NOW, sa više predajnika i sa više prijemnika [13]. ESP-NOW se koristi ukoliko imamo 2 ili više ESP32 uređaja, između kojih želimo da ostvarimo komunikaciju preko Wi-Fi-ja. Potrebno je istaći da je struktura podataka na predajniku i na prijemniku mora biti identična. Ukoliko se odlučimo da koristimo enkriptovanu komunikaciju, LMK mora biti identičan i na prijemnicima i na predajnicima. Takođe, za Wi-Fi komunikaciju možemo izabrati kanale od 1 do 14. Paketi podataka od 250 bajta se šalju i primaju desetinama puta u sekundi, što ih čini pogodnim za podatke koji se brzo menjaju. Navedeni podaci se dobijaju sa uređaja koji vrše akviziciju podataka, kao što su merenje pozicije u prostoru i ugaonog ubrzanja.

V. Dodeljivanje lokalne fiksne IP adrese uređaju

Svaki put kada se uređaj poveže u kućnu ili industrijsku Wi-Fi mrežu, dobija različitu lokalnu IP adresu. Za to je odgovoran DHCP (Dynamic Host Configuration Protocol), što je u potpunosti u redu, ukoliko programer koristi uređaj, ali krajnji korisnik to ne želi. Iz navedenog razloga uređaju dodeljujemo fiksnu lokalnu IP adresu, koja se uklapa u parametre naše kućne Wi-Fi mreže. Na slici 3 je prikazana konfiguracija koju treba da primenimo.

```
IPAddress local_IP(192, 168, 1, 99);
IPAddress gateway(192, 168, 1, 1);
IPAddress subnet(255, 255, 0, 0);
IPAddress primaryDNS(8, 8, 8, 8);
IPAddress secondaryDNS(8, 8, 4, 4);
```

Sl. 3. Postavljanje lokalne fiksne IP adrese i Google DNS (Domain Network Server) servera.

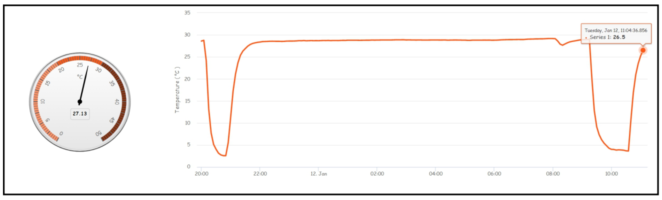
VI. Istraživanje SPIFFS tehnologije

Biblioteka koja omogućava kreiranje kućnog REST (Representational State Transfer) servera je dostupna na Github-u [14].

U Arduino IDE, sve dodatne biblioteke se preuzimaju na sledeći način: "Tools, Manage Libraries, search i install". Ključne reči su: "esp32 HTTPS", a zatim je potrebno instalirati "ESP32_HTTPS_SERVER" biblioteku. Iz navedene biblioteke, koristimo primer "REST-API". Pomoću ovog primera se generiše i Self Signed sertifikat, koji je koristan prilikom razvoja i omogućava upotrebu HTTPS-a. Fajlove sa ekstenzijama: .html, .js, .css i ostale, je potrebno smestiti u folder: "REST-API\data\public". Sve može biti postavljeno na SPIFFS, koristeći opciju iz Tools sekcije: "ESP32 Sketch Data Upload".

VII. Kreiranje web interfejsa IIoT uređaja

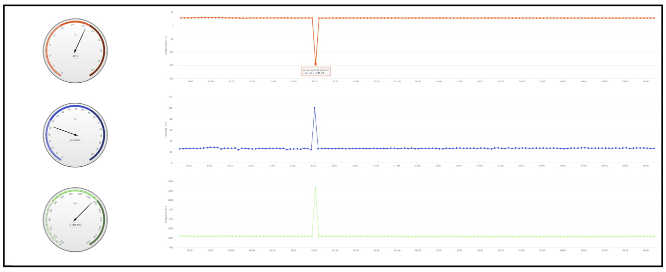
Slika 4 prikazuje deo prilagodljivog web interfejsa sa Highcharts bibliotekom [15], koja je zasnovana na JavaScript-u.



Sl. 4. Deo prilagodljivog interfejsa koji se sastoji od analogne skale i grafika.

VIII. Upravljanje greškama i prevazilaženje problema

Prilikom kreiranja kompleksnih sistema, konstantno vodimo računa o desetinama komponenata. Pored sve pažnje, ponekad se desi da se zaborave osnovni principi, koji su stari koliko je staro i programiranje. Jedan od tih principa je žargonski rečeno, upravljanje "greškama" ili u preciznoj programerskoj terminologiji, upravljanje "izuzecima". Koji god termin koristili, svakako ne smemo da dozvolimo da do korisnika stigne pogrešna ili nepotpuna informacija. Na slici 5 je prikazan web interfejs u kom je zabeležen "loš" signal, koji je direktno plasiran sa hardvera.



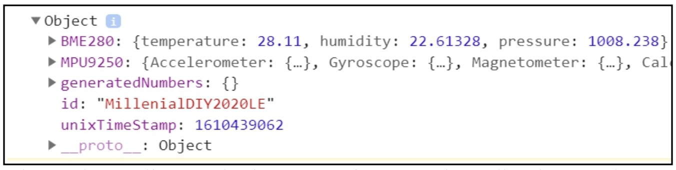
Sl. 5. Web interfejs nakon primanja "lošeg" signala sa hardvera.

Tehnički gledano, u ovom slučaju nije reč o grešci, ali je očigledno da je reč o brojevima koji su izašli van opsega. Korisnik uređaja i/ili aplikacije, to ne želi da vidi. Ukoliko postoji verovatnoća da će se to dogoditi, to mora biti na neki način iskontrolisano, a korisniku treba da bude omogućen kontinualan ispravan rad uređaja, tj. sitema za akviziciju. Navedeni primer predstavlja podsetnik, da moramo biti svesni opsega brojeva koje očekujemo u svakoj komponenti kompleksnog sistema. Treba uzeti u obzir da prilikom očitavanja vrednosti sa senzora Bosch BME280 [16], koji omogućava merenje parametara okruženja, uključujući i atmosferski pritisak, vrednosti budu u opsegu od 300 hPa do 1100 hPa. Sve što je van navedenog opsega je rezultat nekog vida greške, npr. nepotrebnog preopterećenja sa hardverske strane, koje može da prouzrokuje povremena (eng. intermittent) očitavanja "loših" podataka sa senzora. Najveći problem kod povremenih grešaka je činjenica, da ne može da se utvrdi kada će se dogoditi i kako će se one izraziti. Takav tip podataka ne sme da dođe do korisnika, pošto je pogrešan, a takođe se ne sme ni sakriti. Korisniku moraju biti pružene precizne instrukcije šta da radi, bez mnogo tehničkih detalja. Tehnički detalji se zapisuju na takav način da programer može jasno da vidi šta se i kada se dogodilo. Frontend ima mnoštvo naprednih opcija za prikazivanje informacija korisniku, kao što su "toast" notifikacije i popunjavanje polja sa specifičnim porukama. Programeri imaju kreativnu slobodu da izaberu if..else if..else ili try...catch...finally blokove. Dobra analiza sprečava prikaz pogrešnih informacija, kao što je atmosferski pritisak u dnevnoj sobi od 1264 hPa. Da bi bili u potpunosti sigurni da raspolažemo sa ažurnim informacijama, preporučljivo je da se koristi vremenska oznaka (eng. timestamp), Unix tipa, ili bilo koja druga. Vreme može biti očitano sa hardverske komponente kao što je DS3231 [17], preko NTP (Network Time Protocol) servera [18], ili na način po izboru čitaoca. Treba napomenuti da je vremenski žig vrlo bitno parče informacije i ako se pravilno koristi, omogućava da se grafici ne ažuriraju ukoliko ne postoje sveži podaci. Navedeni pristup sprečava širenje dezinformacija.

IX. Test development

Odnos programiranja i testa mora biti bar 1:5. Ukoliko će uređaj imati industrijsku primenu, potrebno je još više testirati. Ekstremne vrednosti mogu biti korisne prilikom testiranja, zato što ako postoji i 1 % šansa da se nešto dogodi, to će se jednom i dogoditi. Kao programeri, moramo biti spremni da izađemo sa tim na kraj i da držimo stvar pod kontrolom. Jedna od prednosti sistema zasnovanom na ESP32, je mogućnost da web interfejs koristimo kao tzv. test bench. Moguće je povezati mnoštvo senzora, dodeliti im ID-je i pratiti njihova očitavanja, npr. na svakih 5 minuta u naredna 24 sata. Mikrokontroler ima dovoljno memorije da prati navedene parametre i bez web interfejsa. Možemo se odlučiti za web interfejs ukoliko želimo da pratimo očitavanja sa mnoštva senzora i da prikazujemo rezultate na uređajima sa velikim i preglednim ekranima, kao što su UHD (Ultra High Definition) televizori rezolucije 3840 px sa 2160 px. Pisanje medija upita (eng. media queries) nije ništa kompleksnije od pisanja medija upita za mobilne telefone i tablete i kreće od : "@media only screen and (max-width: 3840px) {}".

U sred pisanja firmware-a i kreiranja web aplikacije, korisno je pratiti "sirove" podatke (eng. raw data). Na slici 6 su prikazani podaci u JSON (JavaScript Object Notation) formatu, koji se ispisuju u konzoli internet pretraživača, koji su primljeni sa hardvera.



Sl. 6. Primer prikaza podataka u JSON formatu, u konzoli web pretraživača.

Još jedna programerska praksa koja je primenjena, je davanje smislenih naziva promenljivim i njihovo adekvatno grupisanje. Vremenska oznaka 1610439062 prikazuje koliko je sekundi prošlo od 1. januara 1970. godine i predstavlja 12. januar 2021. godine u 08:11:02.

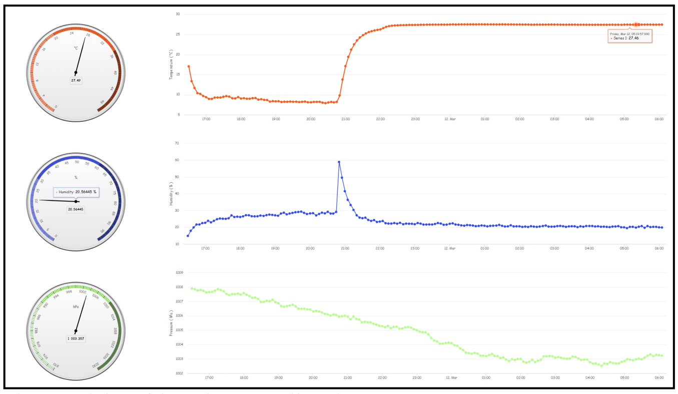
X. Biznis analiza

Mladi inženjeri umeju često da pomešaju biznis analizu sa biznis planom. Biznis analiza nije biznis plan. Biznis analiza je oblast koja predstavlja integralni deo kreiranja bilo kakvog uređaja ili proizvoda, koji će krajnji korisnik koristiti.

Podjednako je kompleksno kao programiranje i projektovanje elektronskih kola, a biznis analitičari najčešće predstavljaju vezu (eng. link) između inženjera i krajnjih korisnika. Na neki način predstavljaju prevodioce, koji zajedno sa korisnicima kreiraju nacrte sistema, a zatim uz maksimalan napor programerima objašnjavaju šta je to što korisnik želi da dobije. Često su biznis analitičari eksperti u njihovim profesijama, koje ne moraju da budu inženjerske.

XI. Regularan web interfejs

Na slici 7 je prikazan web interfejs na kom se vidi znatna promena temperature, a po očitavanju vlažnosti vazduha, na srednjem delu slike se može primetiti da je uređaj premešten iz hladnije prostorije u topliju.



Sl. 7. Web interfejs na kom se prikazuje znatna promena temperature.

Tri vrednosti koje se prikazuju se dobijaju sa istog senzora, Bosch BME280 i nema smisla slati 3 zahteva kada se proverava trenutna vrednost, već je dovoljno to uraditi jednom.

Zahvalnica

Rad su podržali Centar za metrologiju, Fakulteta tehničkih nauka i projekti KALCEA "Knowledge Triangle for a Low Carbon Economy" 618109-EPP-1-2020-1-EL-EPPKA2-CBHE-JP i "Razvoj naučno-stručnih metoda u oblasti metrologije, industrijsko-tehničkih merenja u digitalnom konceptu Industrije 4.0, biomedicinskih mernih sistema i kognitivnih neuronauka primenom napredne metodologije i digitalne tehnologije". Svako pitanje koje student postavi, otvara niz novih mogućnosti i dobar deo ovog rada je nastao upravo kao odgovor na najčešće postavljana pitanja.

Zaključak

Ovaj rad je namenjen svim ambicioznim ljudima, koji ulaze u svet kreiranja uređaja, kao i u programiranje. Takođe, namenjen je i iskusnim programerima i hobistima. Ovaj rad je dokaz koncepta, da je uz entuzijazam moguće krenuti sa izradom industrijskog uređaja, po ceni nižoj od 10$. Ove činjenice omogućavaju korišćenje ovog hardvera i u zemljama u razvoju. Po mišljenju autora, Highcharts se vremenom nametnuo kao vredna zamena Chart.js-u [19], a da je Apache ECharts [20] nešto što bi moglo u skorijoj budućnosti da postane zanimljivo.

Literatura

[1] https://docs.espressif.com/projects/esp-idf/en/latest/esp32/
[2] https://docs.espressif.com/projects/esp-idf/en/latest/esp32/api-reference/network/esp_now.html
[3] https://www.arduino.cc/en/software
[4] https://dl.espressif.com/dl/package_esp32_index.json
[5] https://github.com/me-no-dev/arduino-esp32fs-plugin/releases/
[6] https://github.com/jeffThompson/DarkArduinoTheme
[7] https://github.com/espressif/arduino-esp32
[8] https://learn.sparkfun.com/tutorials/how-to-install-ch340-drivers/all
[9] https://www.silabs.com/developers/usb-to-uart-bridge-vcp-drivers
[10] https://eu.mouser.com/ProductDetail/Espressif-Systems/ESP32-DevKitC-32D?qs=%252BEew9%252B0nqrDsObWEpDx6YQ==
[11] https://eu.mouser.com/ProductDetail/Espressif-Systems/ESP32-DevKitC-32U/?qs=%252BEew9%252B0nqrCEVvpkdH%2FG5Q%3D%3D
[12] https://www.aliexpress.com/item/4001136108709.html?spm=a2g0s.9042311.0.0.7a9b4c4dMRyFWh
[13] https://randomnerdtutorials.com/esp-now-esp32-arduino-ide/
[14] https://github.com/fhessel/esp32_https_server
[15] https://www.highcharts.com/
[16] https://www.bosch-sensortec.com/products/environmental-sensors/humidity-sensors-bme280/
[17] https://www.maximintegrated.com/en/products/analog/real-time-clocks/DS3231.html
[18] https://www.pool.ntp.org/zone/rs
[19] https://www.chartjs.org/
[20] https://echarts.apache.org/en/index.html

Abstract

Abstract: Prior to IIoT (The Industrial Internet of Things), embedded programming and frontend development didn't even found themselves in the same sentence. Hardware that enables seamless integration of these two vast areas is the Espressif ESP32 MCU (MicroController Unit). Once numerous devices were connected to the internet, the security of data became the focal point. ESP-NOW technology is Espressif's proprietary solution for wireless data transfer. It can be encrypted and it enables secure communication between multiple ESP32's. HTTPS (Hypertext Transfer Protocol Secure) is a protocol that increases security on the internet. This article provides a guide for configuring a development environment for ESP32, ESP-NOW example, HTTPS server example, and programming practice intended for error handling. In addition, it offers a responsive example of a graphical user interface of the IIoT device. At the moment, there are numerous possible paths in embedded and frontend programming. This paper follows the path using ESP32 as the hardware, and C++ for the firmware. JavaScript, HTML5, and CSS3 are unavoidable parts of modern industrial devices. Hence, the article provides an example of using JavaScript Highcharts library. This combination of hardware and software is the solution that costs less than 10$, which makes it acceptable in countries under development. Highcharts library is proprietary, and for educational purposes, it is used under Creative Commons (CC) Attribution-Non-Commercial license.

An educational approach to an encrypted data transfer in an embedded and frontend development environment

Ivan Gutai, Platon Sovilj, Marina Subotin, Marjan Urekar, Jelena Milojević, Milica Mitrović

Edukativni primer generisanja i obrade podataka uz alate dostupne u .NET 5, u domenu Metrologije

Ivan Gutai, Member, IEEE, Prof. dr Platon Sovilj, Member, IEEE, Marina Subotin, Member, IEEE, Đorđe Novaković, Member, IEEE, Nemanja Gazivoda, Member, IEEE, Bojan Vujičić, Member, IEEE

Apstrakt— **Baze podataka predstavljaju oslonac kompletne aplikacije, a kreiranje arhitekture i izbor adekvatnih tehnologija iz godine u godinu postaje kompleksnije. Ovaj rad je kreiran kao praktično uputstvo i oslanja se na najnoviju iteraciju popularnog programskog jezika C# 9 i na relacionu bazu podataka, Microsoft-ov SQL Server, skraćeno MSSQL. Entity Framework Core je tehnologija koja značajno olakšava kreiranje i rad sa bazom podataka. Primenjen je eng. code first pristup i iz C# koda je kreirana baza podataka sa dve povezane tabele i inicijalno je popunjena tabela koja predstavlja šifarnik. Nakon kreiranja tabela, iz SQL Server Management Studio (SSMS) su kreirane dve skripte koje omogućavaju kreiranje tabela, bez upotrebe C# koda, tj. napisane su kao Structured Query Language (SQL) skripte. Kreirana je C# aplikacija koja generiše milion zapisa, koji zauzimaju 100 MB u relacionoj bazi i 152 MB kada se serijalizuju u JSON fajl uz pomoć Json.NET-a. Dat je primer korišćenja Xunit alata za testiranje. Dati su primeri proširivanja LINQ (Language-Integrated Query) funkcija Median i Mean. Prikazana su dva načina za očitavanje vrednosti iz baze, uz pomoć LINQ-a i uz pomoć SQL-a.**

Ključne reči— **RDBMS; T-SQL; Entity Framework Core; LINQ; TDD; XUnit; Json.NET; C# 9; .NET 5; Metrologija;**

I. Uvod

Izabrana tehnologija je .NET 5, koja daje naprednije mogućnosti za pisanje višeplatformskih aplikacija. Zbog izuzetne širine, preporučljivo je prvo pročitati [1] i [2], u kojima su detaljno opisane novine koje su unete u programski jezik C# krajem 2020. godine, ali i mnogi osnovni algoritmi. U ovom radu je akcenat stavljen na upotrebu tehnologija za projektovanje i razvoj sistema koji nam omogućava da pisanjem SQL upita ili programski dođemo do rezultata, da ih filtriramo i na kraju sortiramo. Istovremena upotreba LINQ-a i SQL-a za manipulaciju podacima iz baze, predstavlja na neki način verifikaciju LINQ upita.

Ivan Gutai – Fakultet tehničkih nauka, Novi Sad, Srbija (e-mail: gutai@uns.ac.rs).
Platon Sovilj – Fakultet tehničkih nauka, Novi Sad, Srbija (e-mail: platon@uns.ac.rs).
Marina Subotin – Fakultet tehničkih nauka, Novi Sad, Srbija (e-mail: marina.bulat@uns.ac.rs).
Đorđe Novaković – Fakultet tehničkih nauka, Novi Sad, Srbija (e-mail: djordjenovakovic@uns.ac.rs).
Nemanja Gazivoda – Fakultet tehničkih nauka, Novi Sad, Srbija (e-mail: nemanjagazivoda@uns.ac.rs).
Bojan Vujičić – Fakultet tehničkih nauka, Novi Sad, Srbija (e-mail: bojanvuj@uns.ac.rs).

II. Izbor integrisanog razvojnog okruženja

Za razvoj .NET aplikacija su u startu potrebni Visual Studio IDE (Integrated Development Environment) [3] i SDK (Software Development Kit) [4]. Visual Studio Community 2019 je alat, čije će mogućnosti kasnije biti proširivane sa paketima iz NuGet Package Manager-a kao što su: EntityFrameworkCore, EntityFrameworkCore Design, EntityFrameworkCore SqlServer, EntityFrameworkCore Tools, NET.Test.Sdk, NETCore TestHost, xunit, xunit.runner.console, xunit.runner.visualstudio i Newtonsoft.Json. Navedeni alat je besplatan, a samo zahteva logovanje korisnika sa svojim Hotmail nalogom u Visual Studio.

III. .NET 5

Nekoliko godina Microsoft potencira tzv. cross platform razvoj i svake godine nam daje alate, u okviru . NET Core-a, koji to olakšavaju. Naravno, .NET pruža i dalje razvoj aplikacija koje će raditi isključivo na Windows operativnom sistemu, kao što su UWP (Universal Windows Platform) aplikacije, ali omogućava i pravljenje servisa, web ili desktop aplikacija, koje će biti dostupne na različitim platformama. Razlika između .NET Framework i .NET Core je što je inicijalno .NET Core mnogo manji i mi sami biramo koju komponentu framework-a želimo da koristimo u aplikaciji koju pravimo. Da ne bi bilo zabune, definisan je .NET Standard, koji npr. u verziji 2.0 označava koje zajedničke funkcionalnosti imaju .NET Core 2.0 i klasični .NET Framework 4.6.1. [5]. Takođe, .NET Core je izlazio u verzijama 1, 2, 3, pa je 4 preskočena, a aktuelna verzija 5, se označava samo kao .NET 5 (bez Core).

IV. Izbor baze podataka

Dva osnovna mesta za skladištenje podataka su RDBMS (Relational Database Management System) i NoSQL baze podataka. Nekoliko popularnijih RDBMS su: Microsoft SQL Server, PostgreSQL, MySQL i SQLite. Nekoliko popularnijih NoSQL su: Microsoft Azure Cosmos DB, Redis, MongoDB i Apache Cassandra. Ukoliko neko želi da pristupa podacima direktno iz baze, logičan izbor je pisanje SQL upita za RDBMS i pisanje GraphQL upita za NoSQL. U daljem tekstu RDBMS je označen kao relaciona baza, a NoSQL kao nerelaciona baza. Priča o tome koji tip baze podataka je bolji, podseća na priču o tome koji je programski jezik bolji. Jednostavno, ta odluka je na osobi koja projektuje sistem, a u ovom radu je izabrana relaciona baza Microsoft

SQL Server 2019 Express [6], a direktan pristup se vrši preko SSMS (SQL Server Management Studio) [7].

V. KREIRANJE RELACIONE BAZE PODATAKA

U nazivu "EDUG\GI_JOE", "EDUG" je naziv servera, a "GI_JOE" je naziv instance. Baza podataka ima naziv "Expo2021". A u njoj se nalaze 3 tabele: BME280Results, Locations i __EFMigrationsHistory, koja služi za praćenje izmena, a "EF" je skraćeno od "Entity Framework". Tabela Locations ima kolone LocationID i LocationName i stavke 1 i 2 za Office i Balcony, respektivno. Mala i praktična optimizacija, da se vrednosti koje se ponavljaju ne bi iznova upisivale u BME280Results tabelu. BME280ResultID, Timestamp, Temperature, RelativeHumidity, Pressure, LocationID predstavljaju nazive kolona u BME280Results tabeli. Ključevi u tabelama su sledeći:

- ✓ BME280Results tabela, PK_BME280Results i FK_BME280Results_Locations_LocationID,
- ✓ Locations tabela, PK_Locations.

Ukratko, tabela BME280Results ima sekundarni ključ u tabeli Locations, a vezani su sa LocationID-jem. Primarni ključevi u navedenim tabelama su njihove ID kolone. PK i FK su skraćenice od "primary" i "secondary" key, respektivno.

Prvi način kreiranja tabela je bio preko Entity Framework-a. Tačnije, napiše se C# kod (code first pristup) i u konzoli (Package Manager Console) upišemo komande: "Add-Migration" sa nazivom npr. "Initial" i "Update-Database". Za razvoj je korišćen konekcioni string: "Data Source=EDUG\GI_JOE;Initial Catalog=Expo2021;Integrated Security=true;". Drugi način je da se sve to uradi direktno iz SQL-a, što zahteva znatno više vremena. Treći način je da izaberemo u koju bazu ćemo dodavati tabele, a zatim da pokrenemo SQL skripte sa slike 1 i slike 2.

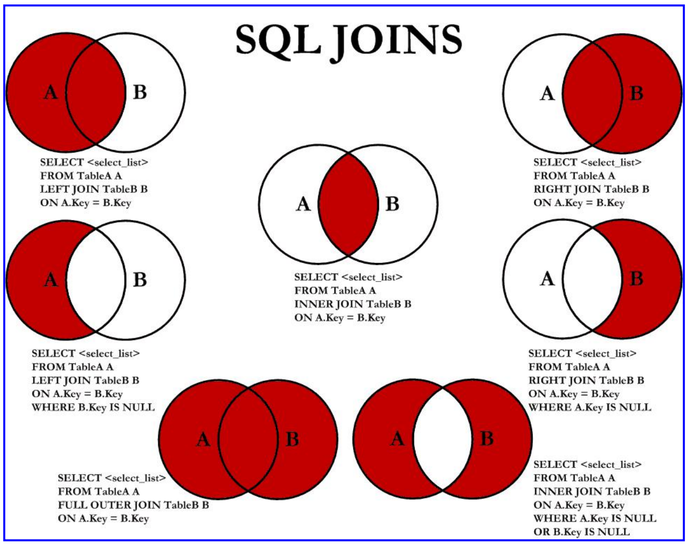
```sql
USE [Expo2021]
GO

/****** Object:  Table [dbo].[Locations]
Script Date: 26.3.2021. 13:16:24 ******/
SET ANSI_NULLS ON
GO

SET QUOTED_IDENTIFIER ON
GO

CREATE TABLE [dbo].[Locations](
    [LocationID] [int] IDENTITY(1,1) NOT NULL,
    [LocationName] [nvarchar](max) NULL,
 CONSTRAINT [PK_Locations] PRIMARY KEY CLUSTERED
(
    [LocationID] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON [PRIMARY]
) ON [PRIMARY] TEXTIMAGE_ON [PRIMARY]
GO
```

Sl. 1. SQL skripta koja omogućava kreiranje tabele Locations, koja je povezana sa tabelom BME280Results.

```sql
USE [Expo2021]
GO

/****** Object:  Table [dbo].[BME280Results]
Script Date: 26.3.2021. 13:16:43 ******/
SET ANSI_NULLS ON
GO

SET QUOTED_IDENTIFIER ON
GO

CREATE TABLE [dbo].[BME280Results](
    [BME280ResultID] [int] IDENTITY(1,1) NOT NULL,
    [Timestamp] [datetime2](7) NOT NULL,
    [Temperature] [money] NOT NULL,
    [RelativeHumidity] [money] NOT NULL,
    [Pressure] [money] NOT NULL,
    [LocationID] [int] NOT NULL,
 CONSTRAINT [PK_BME280Results] PRIMARY KEY CLUSTERED
(
    [BME280ResultID] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON [PRIMARY]
) ON [PRIMARY]
GO

ALTER TABLE [dbo].[BME280Results]
ADD  DEFAULT (getdate()) FOR [Timestamp]
GO

ALTER TABLE [dbo].[BME280Results]  WITH CHECK
ADD  CONSTRAINT [FK_BME280Results_Locations_LocationID]
FOREIGN KEY([LocationID])
REFERENCES [dbo].[Locations] ([LocationID])
ON DELETE CASCADE
GO

ALTER TABLE [dbo].[BME280Results]
CHECK CONSTRAINT [FK_BME280Results_Locations_LocationID]
GO
```

Sl. 2. SQL skripta koja omogućava kreiranje tabele BME280Results, koja je povezana sa tabelom Locations.

Bitan je redosled kojim se navedene skripte pozivaju, zbog međusobne zavisnosti ključeva. Interesantna je i opcija kreiranje View-a u SQL-u, koji na prvi pogled liči na tabelu, a u stvari prikazuje sadržaj iz jedne ili više tabela, u zavisnosti kako to definišemo. Praktičan podsetnik koji detaljnije objašnjava pisanje upita za tzv. spajanje tabela [8] je prikazan na slici 3.



Sl. 3. Vizuelni prikaz razlika između različitih JOIN funkcija.

VI. ENTITY FRAMEWORK CORE TEHNOLOGIJA

Ko se opredeli za .NET tehnologiju, jedan od benefita je Entity framework Core tehnologija, koja omogućava da se upotrebom C# manipuliše bazom podataka, od kreiranja, preko menjanja kolona, do rada sa sadržajem. Na slici 4 je prikazan C# kod koji definiše imena tabela, nazive i tipove kolona koji su kreirani u bazi, a "[Column(TypeName = "money")]" predstavlja jedan od atributa anotacije u Entity Framework Core-u.

```
public class Location
{
    [Key]
    public int LocationID { get; set; }
    public string LocationName { get; set; }
    public virtual ICollection<BME280Result> BME280Results { get; set; }
    public Location()
    {
        this.BME280Results = new List<BME280Result>();
    }
    public class BME280Result
    {
        [DatabaseGenerated(DatabaseGeneratedOption.Identity)]
        public int? BME280ResultID { get; set; }
        [DatabaseGenerated(DatabaseGeneratedOption.Identity)]
        public DateTime Timestamp { get; set; }
        [Column(TypeName = "money")]
        public decimal Temperature { get; set; }
        [Column(TypeName = "money")]
        public decimal RelativeHumidity { get; set; }
        [Column(TypeName = "money")]
        public decimal Pressure { get; set; }
        public int LocationID { get; set; }
        public virtual Location Location { get; set; }
    }
}
```

Sl. 4. C# kod koji definiše imena tabela, nazive i tipove kolona koji su kreirani u bazi.

VII. Generisanje i učitavanje podataka upotrebom C# programskog jezika

Na slikama 5 i 6 su prikazane funkcije koje omogućavaju manipulaciju nad podacima u bazi.

```
public void GenerateNItems(int NoOfItems)
{
    using (var db = new DatabaseContext())
    {
        for (int i = 1; i < NoOfItems + 1; i++)
        {
            var newEntry = new BME280Result();
            //newEntry.BME280ResultID = i;
            newEntry.Temperature = GetRandomNumberInRange(23, 32);
            newEntry.RelativeHumidity = GetRandomNumberInRange(40, 60);
            newEntry.Pressure = GetRandomNumberInRange(1010, 1030);
            newEntry.LocationID = (int)GetRandomNumberInRange(1, 3);
            db.BME280Results.Add(newEntry);

            if (i % 10000 == 0)
            {
                var count = db.SaveChanges();
                Console.WriteLine("{0} records saved to database", count);
            }
            else if (i == NoOfItems)
            {
                var count = db.SaveChanges();
                Console.WriteLine("{0} records saved to database", count);

            }
            //count = db.SaveChanges();
        }
        //Console.WriteLine("{0} records saved to database", count);
    }
}
```

Sl. 5. C# kod koji predstavlja funkciju za kreiranje n stavki u bazi.

```
public void LoadNItems(int NoOfItems)
{
    using (var db = new DatabaseContext())
    {
        Console.WriteLine("{0,16} | {1,-27} | {2,13:N2} | {3,18:N2}" +
            " | {4,10:N2} | {5,14:N2} |", "[BME280ResultID]",
            "[Timestamp]", "[Temperature]", "[RelativeHumidity]",
            "[Pressure]", "[LocationName]");

        var data = db.BME280Results.Join(
        inner: db.Locations,
        outerKeySelector: bm => bm.LocationID,
        innerKeySelector: lo => lo.LocationID,
        resultSelector: (b, l) =>
        new { b.BME280ResultID, b.Timestamp, b.Temperature,
            b.RelativeHumidity, b.Pressure, l.LocationName }).Take(NoOfItems);

        foreach (var item in data)
        {
            Console.WriteLine("{0,16} | {1,27:dd.MM.yyyy. H:mm:ss zzz}" +
                " | {2,-13:N2} | {3,-18:N2} | {4,-10:N2} | {5,14} |"
                , item.BME280ResultID, item.Timestamp, item.Temperature,
                item.RelativeHumidity, item.Pressure, item.LocationName);
        }

    }
}
```

Sl. 6. C# kod koji predstavlja funkciju za učitavanje n stavki iz baze uz upotrebu LINQ-a.

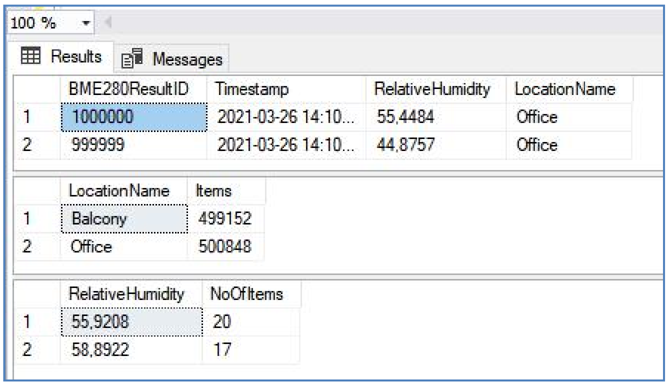
VIII. Osnovni SQL upiti

Na slikama 7 i 8 su prikazani SQL upiti i rezultati, koji omogućavaju sledeće:

- ✓ Selektovanje i kombinovanje rezultata iz dve tabele, od kojih je jedna šifarnik.
- ✓ Selektovanje i grupisanje rezultata u odnosu na lokacije sa kojih su preuzeti podaci.
- ✓ Prikaz modusa iz određene kolone, npr. RelativeHumidity.
- ✓ Funkcija TRUNCATE koju treba koristiti isključivo prilikom testiranja i razvoja. Ova funkcija momentalno briše sve stavke i brojač ID-a vraća na inicijalnu vrednost. Npr. Ako se obrišu sve stavke na ovaj način, prvi sledeći ID koji će biti dodeljen je podrazumevani, u ovom slučaju 1.
- ✓ Funkcija DELETE briše stavke po određenom kriterijumu, a njenom upotrebom se ne gubi zapis o ID-ju stavki. Npr. Ako se obrišu sve stavke na ovaj način, prvi sledeći ID koji će biti dodeljen je 1000001.

```
SELECT TOP (2) B.BME280ResultID, B.Timestamp,
B.RelativeHumidity, L.LocationName
FROM [Expo2021].[dbo].[BME280Results] AS B
INNER JOIN [Expo2021].[dbo].[Locations] AS L
ON B.LocationID = L.LocationID
ORDER BY BME280ResultID DESC

SELECT L.LocationName, COUNT(*) AS Items
FROM [Expo2021].[dbo].[BME280Results] AS B
INNER JOIN [Expo2021].[dbo].[Locations] AS L
ON B.LocationID = L.LocationID
GROUP BY L.LocationName;

SELECT TOP (2) B.RelativeHumidity, COUNT(*) AS NoOfItems
FROM [Expo2021].[dbo].[BME280Results] AS B
INNER JOIN [Expo2021].[dbo].[Locations] AS L
ON B.LocationID = L.LocationID
GROUP BY B.RelativeHumidity
ORDER BY NoOfItems DESC

-- TRUNCATE TABLE [Expo2021].[dbo].[BME280Results]

/*
    DELETE FROM [Expo2021].[dbo].[BME280Results]
    WHERE BME280ResultID > 1000000
*/
```

Sl. 7. SQL upiti koji omogućuju: izlistavanje rezultata, grupisanje rezultata po lokacijama, prikaz modusa za izabranu velilčinu i brisanje podataka.



	BME280ResultID	Timestamp	RelativeHumidity	LocationName
1	1000000	2021-03-26 14:10...	55,4484	Office
2	999999	2021-03-26 14:10...	44,8757	Office

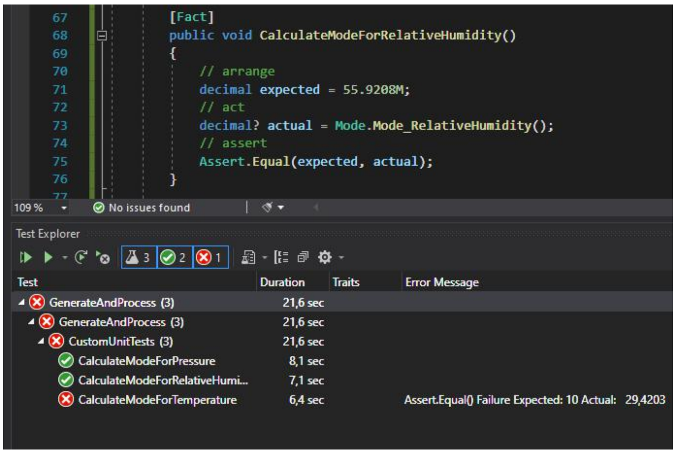
	LocationName	Items
1	Balcony	499152
2	Office	500848

	RelativeHumidity	NoOfItems
1	55,9208	20
2	58,8922	17

Sl. 8. Rezultati SQL upita koji prikazuju TOP n stavki, agregaciju podataka sa dve lokacije i modusa.

IX. Programiranje vođeno testom

TDD (Test driven development) i Unit testovi su postali deo žargona u IT industriji. Najprostiji opis je pisanje koda koji testira kod i daje rezultate o njegovoj ispravnosti. Zasniva se na "AAA principu", tj. "Arrange", "Act" i "Assert", kako god se test framework zvao, a u ovom slučaju je to xUnit [9]. U delu "Arrange" pišemo pretpostavku, u "Act" testiramo funkciju, a u zavisnosti od rezultata, u "Assert"-u je određeno da li je test uspešan ili ne. Interesantno je da ukoliko test "padne", opisano je zbog čega se to desilo. TDD je praktičan na velikim projektima i olakšava regresione testove, zato što automatski proverava da li je nova funkcionalnost narušila neku staru. Na slici 9 je prikazan kod jednog testa i rezultati svih testova u Test explorer-u.



Sl. 9. C# kod jednog testa i rezultati svih testova u Test explorer-u.

Testove treba pisati tako da testiraju stvarne funkcionalnosti, a ne da povećavamo broj testova koji uspešno prolaze. U xUnit-u test je označen sa "[Fact]".

X. Proširivanje ugrađenih LINQ funkcija

Osnovna ideja matematičke statistike jeste da se istraživanja sprovedu na uzorku i da se na osnovu njih donesu zaključci koji se proširuju na populaciju. Statistika koju poseduje LINQ je srednja vrednost (average), a statistike koje su dodate jesu medijana (median) i modus (mean), slika 10.

```
namespace System.Linq
{
    public static class MyLinqExtensions
    {
        // this is a chainable LINQ extension method
        public static IEnumerable<T> ProcessSequence<T>(this IEnumerable<T> sequence)
        {
            // you could do some processing here
            return sequence;
        }
        integers
        public static decimal? Median(this IEnumerable<decimal?> sequence)
        {
            int numberCount = sequence.Count();
            int halfIndex = sequence.Count() / 2;
            var sortedNumbers = sequence.OrderBy(n => n);
            decimal median;
            if ((numberCount % 2) == 0)
            {
                return median = Convert.ToDecimal(((sortedNumbers.ElementAt(halfIndex)
                    + sortedNumbers.ElementAt((halfIndex - 1))) / 2));
            }
            else
            {
                return median = Convert.ToDecimal(sortedNumbers.ElementAt(halfIndex));
            }
        }

        public static decimal? Median<T>(this IEnumerable<T> sequence, Func<T, decimal?> selector)
        {
            return sequence.Select(selector).Median();
        }

        public static decimal? Mode(this IEnumerable<decimal?> sequence)
        {
            var mode = sequence.GroupBy(n => n)
                .OrderByDescending(g => g.Count()).Select(g => g.Key).FirstOrDefault();
            return mode;
        }

        public static decimal? Mode<T>(this IEnumerable<T> sequence, Func<T, decimal?> selector)
        {
            return sequence.Select(selector).Mode();
        }
    }
}
```

Sl. 10. C# kod koji služi za proširivanje funkcija medijane i modusa.

Na slici 11 su prikazane srednje vrednosti, medijane i modusi za 1000000 zapisanih vrednosti, koje sadrže informacije o temperaturi, relativnoj vlažnosti vazduha i pritisku.

```
Mean Temperature: 27,5015
Mean RelativeHumidity: 50,0046
Mean Pressure: 1.020,0059
Median Temperature: 27,5001
Median RelativeHumidity: 50,0038
Median Pressure: 1.020,0053
Mode Temperature: 29,4203
Mode RelativeHumidity: 55,9208
Mode Pressure: 1.018,8240
```

Sl. 11. Rezultati LINQ upita za Average, Mean i Median

Poređenjem rezultata sa slika 8 i 11 se može primetiti da je i LINQ upit ispravno napisan i da rezultati odgovaraju rezultatima dobijenim preko SQL upita.

XI. Primer serijalizacije podataka

Svaki put kada se spominju podaci, na backend-u se često priča o: boxing, unboxing, serialization i deserialization. Prva dva se odnose na objekat, a druga dva na "spuštanje" ili uzimanje objekta iz npr. fajla. Na slikama 12 i 13 je prikazana funkcija koja omogućava serijalizaciju podataka u zadatom obliku, a zatim su prikazani rezultati sa dve stavke, pošto nije baš jednostavno ni na modernom PC-ju otvoriti json fajl od 152 MB, sa svih 1000000 zapisa.

```csharp
public IQueryable GenerateJSON(int NoOfItems)
{
    using (var db = new DatabaseContext())
    {
        var data = db.BME280Results.Join(
        inner: db.Locations,
        outerKeySelector: bm => bm.LocationID,
        innerKeySelector: lo => lo.LocationID,
        resultSelector: (b, l) =>
        new { b.BME280ResultID, b.Timestamp, b.Temperature, b.RelativeHumidity,
            b.Pressure, l.LocationName }).Take(NoOfItems).OrderByDescending(b => b.BME280ResultID);

        var results = data.ToList();

        JObject json =
            new JObject(
                new JProperty("id", "MillenialDIY2020LE"),
                new JProperty("BME280",
                    new JArray(
                            from p in results
                            orderby p.Timestamp
                            select new JObject(
                            new JProperty("BME280ResultID", p.BME280ResultID),
                            new JProperty("Timestamp", p.Timestamp),
                            new JProperty("Temperature", p.Temperature),
                            new JProperty("RelativeHumidity", p.RelativeHumidity),
                            new JProperty("Location", p.LocationName),
                            new JProperty("Pressure", p.Pressure)))));

        File.WriteAllText(@"c:\results.json", JsonConvert.SerializeObject(json));

        // serialize JSON directly to a file
        using (StreamWriter file = File.CreateText(@"c:\results.json"))
        {
            JsonSerializer serializer = new JsonSerializer();
            serializer.Serialize(file, json);
        }

        return data;
    }
}
```

Sl. 12. C# kod koji omogućava da se pomoću LINQ upita izvezu podaci za Average, Mean i Median.

```json
{
    "id": "MillenialDIY2020LE",
    "BME280": [
        {
            "BME280ResultID": 1000001,
            "Timestamp": "2021-03-30T10:41:23.0433333",
            "Temperature": 28.9234,
            "RelativeHumidity": 57.7886,
            "Location": "Office",
            "Pressure": 1020.5552
        },
        {
            "BME280ResultID": 1000002,
            "Timestamp": "2021-03-30T10:41:23.0833333",
            "Temperature": 27.1621,
            "RelativeHumidity": 47.7860,
            "Location": "Office",
            "Pressure": 1019.2326
        }
    ]
}
```

Sl. 13. Sadržaj json fajla.

Zahvalnica

Rad su podržali Centar za metrologiju, Fakulteta tehničkih nauka i projekti KALCEA "Knowledge Triangle for a Low Carbon Economy" 618109-EPP-1-2020-1-EL-EPPKA2- CBHE-JP i "Razvoj naučno-stručnih metoda u oblasti metrologije, industrijsko-tehničkih merenja u digitalnom konceptu Industrije 4.0, biomedicinskih mernih sistema i kognitivnih neuronauka primenom napredne metodologije i digitalne tehnologije". Hvala Microsoft-u što na godišnjem nivou daje nove alate koji olakšavaju rad programerima.

Zaključak

Iako je programiranje podeljeno na backend i frontend, tzv. "fullstack development" omogućava korišćenje najboljih stvari iz oba domena, a to su pisanje algoritama i optimizovanje, kao i smišljanje korisničkog interfejsa. Dobre stvari koje postoje u programiranju ne nestaju, već jednostavno menjaju ime. Npr. Microsoft-ova tehnologija Silverlight i WPF (Windows Presentation Foundation) aplikacije jesu sad stvar prošlosti, ali se XAML (Extensible Application Markup Language), koji je bio sastavna komponenta, jednostavno nastavio koristiti i dalje. Sada se XAML koristi u Xamarin-u, koji omogućava pisanje aplikacija namenjenih za Android operativni sistem. Sada je aktuelna verzija C# 9, a interesantan zapis brojeva je omogućen od verzije 8, npr. da se milion iskaže kao 1_000_000. Još jedan pokazatelj da programski jezici postaju napredniji, je dodavanje podrazumevanog metoda u interfejs, što je u ranijim verzijama C# bilo nezamislivo. U vremenu kada se pojavljuje mnogo novih tehnologija, ko savlada bilo koji "fullstack" set tehnologija, dobija mogućnost efektivnijeg prilagođavanja na novine koje se dešavaju.

Literatura

[1] C# 9 and .NET 5 – Modern Cross-Platform Development: Build intelligent apps, websites, and services with Blazor, ASP.NET Core, and Entity Framework Core using Visual Studio Code, 5th Edition ISBN-13 : 978-1800568105
[2] Fundamentals of Computer Programming with C#: Programming Principles, Object-Oriented Programming, Data Structures (free programming books) ISBN-13 : 978-9544007737
[3] https://visualstudio.microsoft.com/
[4] https://dotnet.microsoft.com/download/visual-studio-sdks
[5] https://docs.microsoft.com/en-us/dotnet/standard/net-standard
[6] https://www.microsoft.com/en-us/sql-server/sql-server-downloads
[7] https://docs.microsoft.com/en-us/sql/ssms/download-sql-server-management-studio-ssms?view=sql-server-ver15
[8] https://clmoffatt.com/visual-sql-joins/
[9] https://xunit.net/

Abstract

Databases represent the backbone of entire applications and creating architecture and choice of adequate technologies annually get more complex. This paper is created as a user manual and uses the newest iteration of popular programming language C# 9 and relational database Microsoft's SQL Server abbreviated MSSQL. Entity Framework Core is a technology that eases the creation and usage of a database. The code first approach is applied, and a database is created from C# code. It has two interconnected tables, and one has initial data and serves as a codebook. After creating tables, from SQL Server Management Studio (SSMS) two scripts are created, without the usage of C#, instead, they are written as Structured Query Language (SQL) scripts. C# application that generates one million records. This takes 100 MB in a relational database and 152 MB when serialized in a JSON file using Json.NET. An example with testing tool Xunit is described. LINQ (Language-Integrated Query) functions Mean and Median are extended. Two ways for getting data from the database are described, using LINQ and using SQL.

AN EDUCATIONAL APPROACH TO GENERATING AND ANALYSING DATA WITH TOOLS AVAILABLE IN .NET 5, IN METROLOGY DOMAIN

Ivan Gutai, Platon Sovilj, Marina Subotin, Đorđe Novaković, Nemanja Gazivoda, Bojan Vujičić

МИКРОЕЛЕКТРОНИКА И ОПТОЕЛЕКТРОНИКА
/
MICROELECTRONICS AND OPTOELECTRONICS, NANOSCIENCES AND NANOTECHNOLOGIES

(MO/MOI)

Efekti zračenja i odžarivanja kod naponsko temperaturno naprezanih p-kanalnih VDMOS tranzistora snage

Sandra Veljković, *Student Member, IEEE,* Nikola Mitrović, *Student Member, IEEE,* Snežana Đorić-Veljković, *Member, IEEE,* Vojkan Davidović, *Member, IEEE,* Snežana Golubović, *Member, IEEE,* Danijel Danković, *Member, IEEE*

***Sadržaj* — U radu su prikazani efekti zračenja i odžarivanja kod p-kanalnih VDMOS (vertical double-diffused metal-oxide semiconductor) tranzistora snage koji su prethodno bili izloženi naponsko temperaturnom naprezanju tokom različitih unapred tačno utvrđenih vremenskih perioda. Ovaj eksperiment je sproveden kako bi se ispitali efekti naprezanja zračenjem kod komponenata koje su prethodno bile izložene drugim uticajima, odnosno naprezanjima. Primećeno je da je promena napona praga tokom ozračivanja malo više izražena kod komponenata koje su bile izložene naponsko temperaturnom naprezanju negativnom polarizacijom gejta nedelju dana, što može biti od značaja u slučaju kada su komponente primile visoke ukupne doze. Pored toga, primećeno je da kod komponenata koje su naponski temperaturno naprezane jedan sat, a zatim ozračene do 90 Gy pri pozitivnoj polarizaciji gejta, termički aktivirani procesi tokom odžarivanja nisu dovoljni da se napon praga smanji na vrednost pre naponsko temperaturnih naprezanja, što bi moglo biti od interesa u slučaju kada komponente rade u pooštrenim uslovima.**

***Ključne reči* — VDMOS tranzistor snage; napon praga; NBT (negative bias temperature) nestabilnosti; ozračivanje; odžarivanje.**

I. Uvod

Zbog svojih specifičnih performansi VDMOS (vertical double-diffused metal-oxide semiconductor) tranzistori snage imaju široku primenu kako u komercijalnim, tako i u uređajima specijalne namene (u prekidačkim izvorima napajanja, u audio pojačavačima, u automobilskoj industriji u uređajima i sistemima koji predstavljaju i dodatnu i primarnu opremu [1-3]). Pri tome, u nekim aplikacijama mogu biti podvrgnuti i pooštrenim uslovima rada i/ili nekom od oblika naprezanja, tako da u poslednjih dvadeset godina postoji povećano interesovanje za istraživanje njihove pouzdanosti, kao i efekata koji se javljaju pri radu u određenim uslovima [2-20]. Istraživanja su bila usredsređena uglavnom na efekte koji se javljaju pri izlaganju visokim vrednostima električnog polja [4-6], pri ozračivanju [2, 7-9], kao i pri ispitivanju efekata ubrzanih visokom temperaturom [10-12] i efekata naponsko temperaturnih naprezanja [4, 13-19].

VDMOS tranzistori snage su veoma osetljivi na jonizujuće zračenje koje ozbiljno degradira napon praga (V_T) [2, 7-9], tako da je za primenu u radijacionom okruženju, postupak provere pouzdanosti važan za odabir visoko pouzdanih komponenata. Standardno ispitivanje pouzdanosti za ove komponente uključuje „burn-in“ testove, pri kojima se primenjuju polarizacija i povišena temperatura [10-12, 21, 22].

Naponsko temperaturne nestabilnosti koje nastaju usled negativne polarizacije gejta i povišene temperature (u literaturi poznate kao NBTI - negative bias temperature instability), su nestabilnosti koje kod p-kanalnog MOSFET-a dovode do degradacije veoma bitnih električnih parametara, od kojih je najkritičniji napon praga. Ove nestabilnosti se javljaju kao rezultat rada uređaja na povišenoj temperaturi (obično 100-250 ^{o}C) i sa negativnom polarizacijom gejta (pri električnom polju u oksidu gejta u opsegu 2-6 MV/cm) [13, 16, 23-25]. Iako se ovi uslovi mogu javiti pri sprovođenju testova [10, 11], mogu da se jave i u nekim primenama pri standardnom režimu rada VDMOS tranzistora snage [13]. Štaviše, ove komponente mogu biti izložene istovremeno i zračenju i NBT naprezanju, kao što je navedeno u [26-29]. Takođe, u nekim primenama je moguća kombinacija sa drugim vrstama naprezanja, poput izlaganja magnetnom polju [18-20] ili uslovima visoke temperature i vlažnosti [10].

Promena napona praga IRF9520 p-kanalnih VDMOS tranzistora snage izazvana NBT naprezanjem (koja se aktivno proučava već više godina [13, 16, 30-32]) se menja po dobro poznatom t^n zakonu. Pri tome se ova zavisnost može da podeli na tri faze u zavisnosti od vrednosti parametra *n*. Vremena koja odgovaraju kraju prve faze (1 sat) i druge faze (1 nedelja, tj. 168 sati) izabrana su kao periodi u toku kojih je naprezanje negativnom polarizacijom gejta primenjeno u eksperimentima čiji su rezultati analizirani u ovom radu.

U ovom radu prikazani su rezultati ozračivanja IRF9520 p-kanalnih VDMOS tranzistora snage, koji su prethodno naponski temperaturno naprezani negativnom polarizacijom gejta (NTNNP) 1 sat i 168 sati, sa ciljem da se pojasni efekat trajanja naprezanja negativnom polarizacijom gejta na kasnije efekte zračenja i odžarivanja. Na osnovu podataka o apsorbovanoj dozi pri tipičnoj primeni MOS komponenata u komunikacionim satelitima [33], izabrana je granica ukupne doze od 100 Gy (SiO_2).

Sandra Veljković – Univerzitet u Nišu, Elektronski fakultet, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: sandra.veljkovic@elfak.rs).

Nikola Mitrović – Univerzitet u Nišu, Elektronski fakultet, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: nikola.i.mitrovic@elfak.ni.ac.rs).

Snežana Đorić-Veljković – Univerzitet u Nišu, Građevinsko-arhitektonski fakultet, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: snezana.djoric.veljkovic@elfak.ni.ac.rs).

Vojkan Davidović – Univerzitet u Nišu, Elektronski fakultet, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: vojkan.davidovic@elfak.ni.ac.rs).

Snežana Golubović – Univerzitet u Nišu, Elektronski fakultet, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: snezana.golubovic@elfak.ni.ac.rs).

Danijel Danković – Univerzitet u Nišu, Elektronski fakultet, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: danijel.dankovic@elfak.ni.ac.rs).

II. EKSPERIMENT

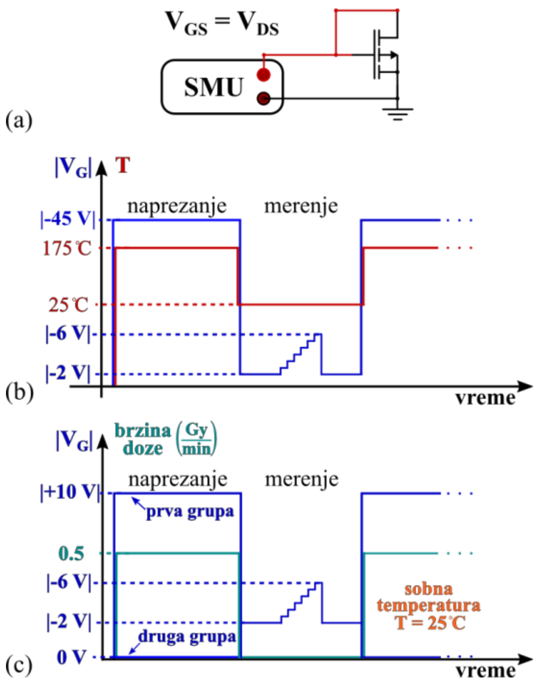
Uzorci koji su testirani u eksperimentu su komercijalni, p-kanalni VDMOS tranzistori snage tipa IRF9520. Eksperimentalna procedura se sastojala od primene NTNNP, izlaganja γ-zračenju i termičkog odžarivanja, koji su međusobno bili odvojeni dvema fazama spontanog oporavka.

Primena NTNNP na komponente je obavljena u Heraeus HEP2 komorama za testiranje pouzdanosti, koje su obezbeđivale stabilnu temperaturu od 175 ºC tokom 168 sati za prvu grupu uzoraka i 1 sat za drugu grupu uzoraka. Kod svih komponenata sors i drejn su bili uzemljeni, dok je primenjena polarizacija na gejtu bila - 45 V.

Uzorci su bili izloženi γ-zračenju korišćenjem izvora Co-60 sa brzinom doze od 0.5 Gy(SiO_2)/min u Laboratoriji za zaštitu od zračenja i zaštitu životne sredine, Instituta za nuklearne nauke "Vinča" u Srbiji. Obe grupe komponenata koje su najpre podvrgnute NTNNP (168 sati i 1 sat), kao i one koje nisu prethodno naprezane podeljene su u po dve podgrupe. Komponente prve podgrupe su ozračivane bez polarizacije, dok su komponente druge podgrupe ozračivane uz primenu pozitivne polarizacije gejta od +10 V, pri čemu su sors i drejn svih komponenata bili uzemljeni. Svih šest podgrupa komponenata su ozračene do ukupne doze od 90 Gy.

Završna faza eksperimenta, termičko odžarivanje, izvedeno je na 175°C tokom 168 sati bez primenjene polarizacije. Pored toga, treba napomenuti da su se komponente spontano oporavljale između NTNNP i zračenja (tokom 670 sati) i između zračenja i odžarivanja (tokom 15 sati), u oba slučaja na sobnoj temperaturi bez polarizacije komponenata.

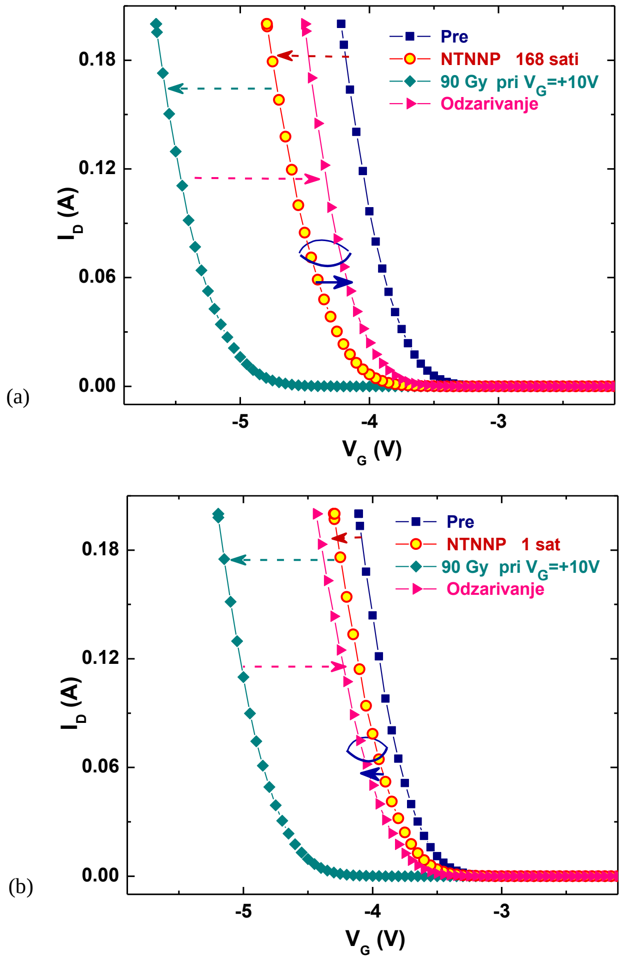
U cilju praćenja degradacije komponenata, NTNNP, ozračivanje i odžarivanje su prekidani nakon prethodno definisanog vremenskog perioda da bi se snimale prenosne karakteristike I_D-V_G, na osnovu kojih je određen napon praga. Sva merenja su obavljana na sobnoj temperaturi. Na Sl. 1 su prikazani merna konfiguracija i uslovi merenja. Za električnu karakterizaciju je korišćen visoko precizni Keysight Technologies B2901A (Source Measure Unit - SMU), koji je kontrolisan pomoću laptopa, preko GPIB-a.



Sl. 1. Merenja istraživanih komponenata a) šema, i uslovi tokom naprezanja b) pri NTNNP i c) pri ozračivanju.

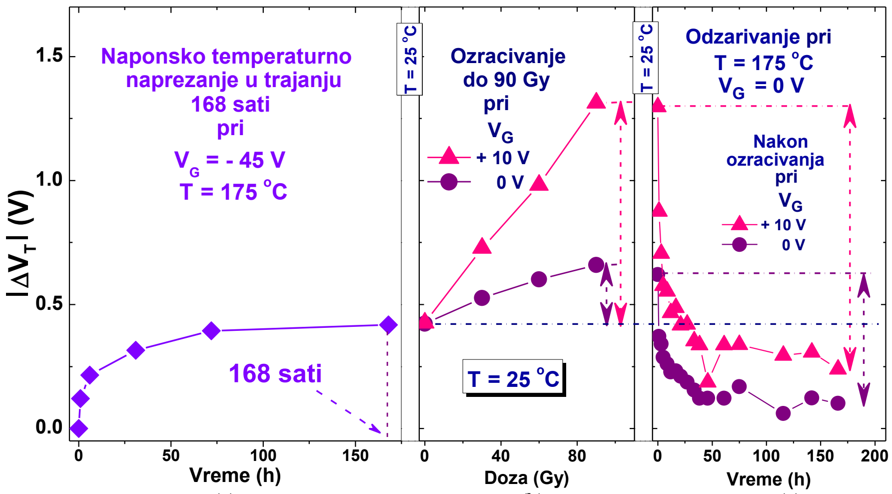
III. REZULTATI I DISKUSIJA

Iako je u toku svakog procesa naprezanja vršena karakterizacija, na Sl. 2 su prikazane samo prenosne strujno naponske karakteristike nakon NTNNP, ozračivanja do 90 Gy (pri V_G = +10 V) i odžarivanja. Pri tome su na Sl. 2a i Sl. 2b predstavljene neke od prenosnih karakteristika p-kanalnih VDMOS tranzistora snage koji su podvrgnuti NTNNP tokom 168 sati i 1-og sata, respektivno.



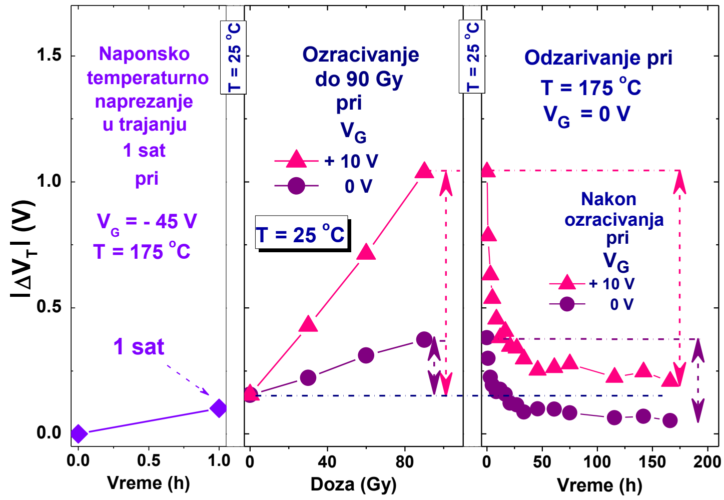
Sl. 2. Prenosne karakteristike p-kanalnog VDMOS tranzistora snage (IRF9520), nakon (a) 168 sati i (b) 1-og sata naponsko temperaturnog naprezanja, ozračivanja pri V_G = + 10 V i odžarivanja.

Na osnovu Sl. 2 se jasno uočava da nakon primene NTNNP dolazi do pomeranja prenosnih karakteristika duž V_G ose ka negativnijim vrednostima, što ukazuje na porast napona praga ispitivanih tranzistora po apsolutnoj vrednosti, pri čemu su ove promene jasno izraženije kod uzoraka koji su duže naprezani, kao što se vidi na Sl. 3a i Sl. 4a. Ozračivanje do 90 Gy, koje je sledilo, je dovelo do značajnog daljeg pomeranja prenosnih karakteristika duž V_G ose ka negativnijim vrednostima, kod obe grupe komponenata, odnosno do porasta V_T po apsolutnoj vrednosti (Sl. 3b i Sl. 4b), što je uočeno i kod komponenata koje prethodno nisu podvrgnute NTNNP, kao što se vidi na Sl. 5a.



(a) (b) (c)

Sl. 3. Promene V_T p-kanalnih VDMOS tranzistora snage (IRF9520) tokom primene (a) NTNNP u trajanju od 168 sati, (b) ozračivanja i (c) odžarivanja.

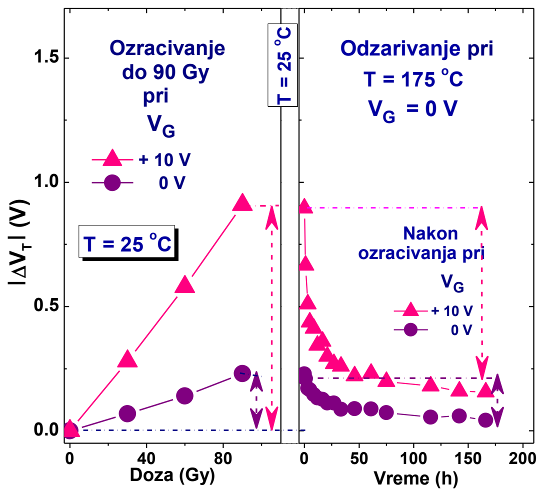


(a) (b) (c)

Sl. 4. Promene V_T p-kanalnih VDMOS tranzistora snage (IRF9520) tokom primene (a) NTNNP u trajanju od 1-og sata, (b) ozračivanja i (c) odžarivanja.

Tokom ozračivanja komponenata kod kojih je primenjena polarizacija gejta došlo je do veće promene V_T usled formiranja znatno većeg broja defekata u oksidu gejta i zahvatanja nosilaca naelektrisanja na površinskim stanjima. Uopšte, dobro je poznato da su promene V_T prouzrokovane formiranjem naelektrisanja u oksidu gejta i površinskih stanja tokom određenog naprezanja usled elektrohemijskih procesa koji uključuju defekte u oksidu, na međupovršini, šupljine i čestice vodonika.

Jonizujuće zračenje formira parove elektron-šupljina u strukturi gejta oksida i mada se neki parovi rekombinuju najveći broj elektrona napusti oksid [26]. Ukoliko je primenjena pozitivna polarizacija na gejt elektroni napuštaju oksid upravo kroz gejt, dok se najveći broj šupljina (koje imaju manju pokretljivost) zahvata na defektima, pri čemu dolazi do povećanja zahvaćenog pozitivnog naelektrisanja u oksidu [28]. U slučaju ozračivanja bez polarizacije gejta, u oksidu postoji malo pozitivno električno polje kao rezultat male kontaktne razlike potencijala između polisilicijumskog gejta i n-tipa balka kod p-kanalog MOSFET-a i može uticati na rezultate formiranja naelektrisanja u oksidu gejta pod dejstvom γ zračenja [34, 35].



(a) (b)

Sl. 5. Promene napona praga p-kanalnih VDMOS tranzistora snage (IRF9520) tokom (a) ozračivanja pri V_G = + 10 V i (b) odžarivanja.

Pored toga, konačni efekti ozračivanja zavise i od predistorije komponenata, tako da su konačne apsolutne vrednosti V_T najviše kod komponenata prethodno podvrgnutih NTNNP u toku 168 sati, dok su najmanje kod komponenata koje prethodno nisu podvrgnute NTNNP. Međutim, brzina promene V_T u toku procesa ozračivanja je nešto manje izražena kod komponenata podvrgnutih NTNNP, što može biti posledica manje brzine porasta i gustine naelektrisanja u oksidu gejta i površinskih stanja zbog smanjenog broja prekursora defekata nakon NTNNP. Pri tome su promene V_T tokom zračenja nešto izraženije kod komponenata prethodno podvrgnutih NTNNP tokom 168 sati nego tokom 1-og sata i ova razlika može biti posledica razlike u nastalim i započetim procesima u prvoj i drugoj fazi primenjenog NTNNP. Srednje vrednosti razlika u promenama napona praga nakon ozračivanja do 90 Gy (ΔV_T (90 Gy)) i nakon primene NTNNP (ΔV_T (NTNNP)) su naznačene strelicama na desnoj strani Sl. 3b, Sl. 4b, a na Sl. 5a samo nakon ozračivanja i date su u Tab. 1.

Tabela 1. Srednje vrednosti promene ΔV_T nakon ozračivanja u odnosu na vrednosti nakon primenjenog NTNNP.

	Vreme primenjenog NTNNP		
	168 sati	1 sat	0 (bez NTNNP)
V_G tokom ozračivanja	ΔV_T (90 Gy) - ΔV_T (NTNNP) (V)		ΔV_T (90 Gy)
+ 10 V	0.897	0.882	0.912
0 V	0.220	0.218	0.232

Za komponente ozračene do 90 Gy, pri V_G=+10 V razlika je 0.16 mV/Gy između komponenata na koje je primenjeno NTNNP 168 sati i 1 sat. Iako ove vrednosti nisu posebno velike mogu biti od interesa u slučaju ozračivanja do visokih doza, koje mogu biti primljene tokom rada uređaja u elektronskim sistema pri dugotrajnoj misiji u radijacionom okruženju.

Odžarivanje koje je primenjeno nakon ozračivanja i kratkotrajnog spontanog oporavka (Sl. 3c, Sl. 4c i Sl. 5b) dovodi do značajnog opadanja V_T po apsolutnim vrednostima. Pri tome je ovo opadanje najveće na samom početku, kod svih grupa komponenata i izraženije je ukoliko je pri ozračivanju primenjena polarizacija gejta. Vrednosti promena V_T u toku odžarivanja, kod komponenata na koje je primenjeno NTNNP (Sl. 3c i Sl. 4c), kao i na koje nije primenjeno NTNNP (Sl. 5b) date su u Tab. 2, a ove promene su naznačene i strelicama na desnoj strani ovih slika.

Treba napomenuti da konačne vrednosti ΔV_T postignute nakon vremena odžarivanja od 168 sati na temperaturi od 175°C, u odnosu na vrednosti posle primene NTNNP, kvalitativno zavise od trajanja primenjenog NTNNP. Naime, kod svih komponenata na koje je primenjeno NTNNP u toku 168 sati, konačne (apsolutne) vrednosti ΔV_T nakon odžarivanja su niže od onih neposredno nakon NTNNP. Međutim, kod komponenata na koje je primenjeno NTNNP u toku jednog sata vrednosti ΔV_T nakon odžarivanja su niže od onih neposredno nakon NTNNP za komponente ozračene bez polarizacije (pri V_G=0 V), dok su za komponente ozračene sa polarizacijom (pri V_G=+10 V) više. Ove razlike u vrednostima promena V_T mogu se videti sa Sl. 3c, i Sl. 4c, a mogu se jasno uočiti i na Sl. 2, gde je naznačen odnos karakteristika nakon NTNNP i nakon odžarivanja.

Tabela 2. Srednje vrednosti promene ΔV_T u toku odžarivanja.

	Vreme primenjenog NTNNP		
	168 sati	1 sat	0 (bez NTNNP)
V_G tokom ozračivanja	ΔV_T u toku odžarivanja (V)		
+ 10 V	1.0582	0.8306	0.7842
0 V	0.5194	0.3290	0.1980

Tokom odžarivanja ispitivane komponente nisu bile polarisane, pa su efekti koji se javljaju u komponentama termički aktivirani, tako da procesi koji se dešavaju i elektrohemijske reakcije mogu biti u vezi sa difuzijom nekih čestica vodonika. Tako čestice kao što su neutralni molekuli H_2 i pozitivni joni H^+ mogu da difunduju sa mesta gde je njihova koncentracija veća (oksid) ka mestima sa nižom koncentracijom (međupovršina), a može doći i do preraspodela zahvaćenog pozitivnog naelektrisanja u oksidu [36, 37]. Očigledno termički aktivirani procesi odžarivanja bez polarizacije (koji su u vezi sa difuzijom čestica vodonika) nisu dovoljni da smanje vrednosti promena napona praga na vrednost pre primene NTNNP kod komponenata na koje je primenjeno NTNNP u toku jednog sata i ozračivanih do 90 Gy pod polarizacijom V_G=+10 V.

IV. Zaključak

U ovom radu su prikazani eksperimentalni rezultati zračenja i odžarivanja p-kanalnih VDMOS tranzistora snage nakon primene naponsko temperaturnog naprezanja negativnom polarizacijom gejta tokom 168 sati i 1 sat. Svrha istraživanja je bila da se rasvetli trajanje NTNNP na efekte zračenja i odžarivanja, kako bi se ispitali efekti specifičnog uticaja zračenja na uređaje koji su prethodno bili izloženi naprezanju. Pokazano je da zračenje dovodi do daljeg porasta V_T kod svih komponenata koje su prethodno izložene NTNNP, pri čemu je ovaj porast malo izraženiji kod komponenata koje su prethodno bile izložene NTNNP tokom 168 sati, nego kod onih koje su bile izložene tokom jednog sata. To bi moglo biti od interesa u slučaju ozračivanja do visokih doza, koje mogu biti primljene tokom rada uređaja u elektronskim sistemima pri dugotrajnoj misiji u radijacionom okruženju. Takođe, primećeno je da kod uređaja izloženih NTNNP tokom jednog sata i ozračenih do 90 Gy pod polarizacijom, termički aktivirani procesi tokom odžarivanja (povezani sa difuzijom neutralnih čestica, kao što su molekuli vodonika) nisu dovoljni za smanjenje V_T na vrednost pre izlaganja NTNNP. To može biti od interesa kada se komponente ugrađuju u uređaje koji rade u pooštrenim uslovima, u okruženja gde su izloženi višestrukim naprezanjima.

ZAHVALNICA

Prikazani rezultati dobijeni su u okviru programa za finansiranje naučnoistraživačkog rada (pod brojem 451-03-9/2021-14/200102) koji finansira Ministarstvo prosvete, nauke i tehnološkog razvoja Republike Srbije i u okviru projekta “Osobine tankih i ultratankih oksidnih slojeva” (F-148), koji finansira Srpska akademija nauka i umetnosti-SANU. Rezultatima predstavljenim u ovom radu je u mnogome doprineo nedavno preminuli akademik Ninoslav Stojadinović i svojim naučnim i istraživačkim aktivnostima i značajnom saradnjom sa autorima.

LITERATURA

[1] B. Jayant Baliga, *Fundamentals of Semiconductor Power Devices*, Springer, New York, 2008.
[2] T. P. Ma and P. V. Dressendorfer, Ionizing Radiation Effects in MOS Devices and Circuits, 1989, Wiley; New York.
[3] S. Gamerith and M. Polzl, Negative bias temperature stress in low voltage p-channel DMOS transistors and role of nitrogen, *Microelectron. Reliab.* 42, 2002, 1439-1443.
[4] N. Stojadinović et al., Effects of electrical stressing in power VDMOSFETs, Microelectron. Reliab. 45, 2005, 115-122.
[5] G. S. Ristić, et al. Defect behaviors in n-channel power VDMOSFETs during HEFS and thermal post-HEFS annealing, Appl. Surf. Sci. 252 (8), 2006, 3023–3032.
[6] X. R. Ye et al., VDMOSFET HEF degradation modelling considering turn-around phenomenon, *Microelectron. Reliab.* 50, 2018, 37-41.
[7] N. Stojadinović et al., Analysis of gamma-irradiation induced degradation mechanisms in power VDMOSFETs, *Microelectron. Reliab.* 35, 1995, 587-602.
[8] C. Pickard et al., Use of commercial VDMOSFETs in electronic systems subjected to radiation, *IEEE Trans. Nucl. Sci.* 47, 2000, 627-633.
[9] M. S. Park, C. R. Wie, Study of radiation effects in γ-ray irradiated power VDMOSFET by DCIV technique, *IEEE Trans. Nucl. Sci.* 48, 2001, 2285-2293.
[10] N. Tošić, et al., High-temperature-reverse-bias testing of power VDMOS transistors, *Microelec. Reliab.* 37, 1997, 1759-1762.
[11] S. Djorić-Veljković et al, Effects of burn-in stressing on post-irradiation annealing response of power VDMOSFETs, *Microelectron. Reliab.* 43, 2003, 1455-1460.
[12] P. Magnone et al., Investigation of degradation mechanisms in low-voltage p-channel power MOSFETs under High Temperature Gate Bias stress, *Microelectron. Reliab.* 88-90, 2018, 438-442.
[13] N. Stojadinović et al., Negative bias temperature instability mechanisms in p-channel power VDMOSFETs, *Microelectron. Reliab.* 45, 2005, 1343-1348.
[14] I. Manić et al., NBTI related degradation and lifetime estimation in p-channel power VDMOSFETs under the static and pulsed NBT stress conditions, *Microelectron. Reliab.* 51, 2011, 1540-1543.
[15] A. N. Tallarico et al., Negative bias temperature stress reliability in trench-gated p-channel power MOSFETs, *IEEE Trans. Dev. Mater. Reliab.* 14, 2014, 657-663.
[16] N. Stojadinović et al., Negative Bias Temperature Instability in Thick Gate Oxides for Power MOS Transistors, in *Bias Temperature Instability for Devices and Circuits*, edited by Tibor Grasser, 2014, Springer, New York, 533-559.
[17] D. Danković et al., Negative bias temperature instability in p-channel power VDMOSFETs: recoverable versus permanent degradation, *Semicond. Sci. Technol.* 30, 2015, 105009.
[18] H. Tahi et al., Experimental Investigation of NBTI Degradation in Power VDMOS Transistors Under Low Magnetic Field, *IEEE Trans. Device Mater. Reliab.* 17 (1), 2017, 99-105.
[19] A. Bhattacharjee, S.N. Pradhan, NBTI-Aware Power Gating Design with Dynamically Varying Stress Probability Control on Sleep Transistor, *J. Circuit Syst. Comp.* 2021, Art. no 2120004 2120004.
[20] H. Tahi et al., Capacitance–Voltage Technique Based on Time Varying Magnetic Field for VDMOSFET-Part I: Concept and Implementation, *IEEE Trans. on Elec. Dev.* 68 (5), 2021, 2173-2180
[21] N. Stojadinović et al., Effects of High Electric Field and Elevated-Temperature Bias Stressing on Radiation Response in Power VDMOSFETs, *Microelectron. Reliab.* 42, 2002, 669-677.
[22] N. Stojadinović et al., Effects of Burn-in Stressing on Radiation Response of Power VDMOSFETs, *Microelectron. J.*, 33, 2002, 899-905.
[23] D. K. Schröder, J. A. Babcock, Negative bias temperature instability: Road to cross in deep submicron silicon semiconductor manufacturing, *J. Appl. Phys.* 94, 2003, 1-18.
[24] J. H. Stathis, S. Zafar, The negative bias temperature instability in MOS devices: A Review, *Microelectron. Reliab.* 46, 2006, 270-286.
[25] T. Grasser et al., Recent advances in understanding the bias temperature instability, *IEEE Proc. IEDM*, 2010, 82-85.
[26] V. Davidović et al., NBTI and irradiation effects in p-channel power VDMOS Transistors, *IEEE Trans. Nucl. Sci.* 63, 2016, 1268-1275.
[27] V. Davidović et al., Effects of consecutive irradiation and bias temperature stress in p-channel power vertical double-diffused metal oxide semiconductor transistors, *Jap. J. App. Phys.* 57 (4), 2018, 044101-1-10.
[28] V. Davidović et al., NBT Stress and Radiation Related Degradation and Underlying Mechanisms in P-Channel Power VDMOS Transistors, *Facta Universitatis: Electronics and Energetics*, 16 (4), 2018, 367-388.
[29] D. Danković, et al., Radiation and Annealing Related Effects in NBT Stressed P-Channel Power VDMOSFETs, prihvaćeno za *32nd European Symposium on Reliability of Electron Devices, Failure Physics and Analysis (ESREF 2021*), Bordeaux (France) 8-8 October 2021
[30] D. Danković et al., Implications of Negative Bias Temperature Instability in Power MOS Transistors, in *Micro Electronic and Mechanical Systems*, edited by Kenichi Takahata, 2009, IN-TECH Press, Boca Raton, 19.319-19.342
[31] D. Danković et al., Negative Bias Temperature Instabilities in Sequentialy Stressed and Annealed in p-Channel Power VDMOSFETs, *Microelectron. Reliab.* 47, 2007, 1400-1405.
[32] I. Manić et al., Effects of Low Gate Bias Annealing in NBT Stressed p-Channel Power VDMOSFETs, *Microelectron. Reliab.* 49, 2009, 1003-1007
[33] D.M. Fleetwood et al., An overview of radiation effects on electronics in the space telecommunication environment, *Microelectron. Reliab*. 40, 2000, 17-26.
[34] T. R. Oldham and J. M. McGarrity, Comparison of ^{60}Co response and 10keV X-ray response in MOS capacitors, *IEEE Trans. Nucl. Sci.*, 30, 1983, 4377-4381.
[35] M.M. Pejovic, Application of p-channel power VDMOSFET as a high radiation doses sensor, *IEEE Trans. Nucl. Sci.* 62, 2015, 1905-1910.
[36] D. M. Fleetwood, Effects of hydrogen transport and reactions on microelectronics radiation response and reliability, *Microelectron. Reliab*. 42, 2002, 523-541.
[37] T. Grasser et al., The paradigm shift in understanding the bias temperature instability: from reaction-diffusion to switching oxide traps, *IEEE Trans. Electron. Dev.* 58, 2011, 3652-3666.

ABSTRACT

The radiation and annealing effects in p-channel VDMOS power transistors which were formerly exposed to the negative bias temperature stress during different previously exactly determined time periods are presented in the paper. This experiment was conducted in order to explore the effects of particular radiation stress in components which were formerly exposed to other influences that may act as stress. It was noticed that radiation response of threshold voltage was slightly further changed for components which were previously NBT (negative bias temperature) stressed for one week, what may be of importance when high total doses were received. Besides, it was noticed that in components exposed to NBT stress for one hour before irradiation (up to 90 Gy) with positive gate voltage applied, thermally activated processes during subsequent annealing are not enough to reduce V_T to the magnitude before NBT stress, which might be of interest in the case when components work under harsh conditions.

Effects of Radiation and Annealing in Bias Temperature Stressed p-channel power VDMOS transistors

SandraVeljković, Nikola Mitrović, Snežana Đorić-Veljković, Vojkan Davidović, Snežana Golubović and Danijel Danković

Porast elektroprovodnosti Li-jonskih baterija oblaganjem elektroda metal-oksidnim nano-filmovima

Jovan P. Šetrajčić, Siniša M. Vučenović, Igor J. Šetrajčić, Stevo K. Jaćimovski, Ana J. Šetrajčić-Tomić, Dušan I. Ilić, Nikola R. Vojnović

***Apstrakt* — U radu su predstavljeni rezultati teoriskih istraživanja ponašanja fononskog podsistema u ultratankim metal-oksidnim (MO) prevlakama kakve se nanose na elektrode u Li-jonskim baterijama radi povećanja efikasnosti jonskog transporta. Primenjen je metod Grinovih funkcija i numeričkom analizom je pokazano da kod ultratankih filmova i to normalno na granične površi tog filma, dolazi do pojave izrazito pojačanog mehaničkog oscilovanja kristalne rešetke i formiranja stojećih talasa, čime fononi podstiču oslobađanje zarobljenih jona u graničnim slojevima elektroda i time povećavaju efikasnost jonskog transporta između elektroda, a time i doprinose uvećanju elektro-provođenja ovih baterija.**

***Ključne reči*—Li-jonske baterije; jonska provodnost; fononi; ultratanke film-strukture.**

I. UVOD

U ovom radu će se analizirati nanostrukturni materijali za poboljšanje energetske konverzije i skladištenje energije. Li-jon baterije bi trebalo da mogu da zadovolje izuzetno visoko postavljene ciljeve jer imaju vrlo visoku zapreminsku (oko 300 Wh/m^3) i težinsku (oko 130 Wh/kg) specifičnu energetsku gustinu [1], tj. imaju visoku kapacitivnost, veliku gustinu energije i radni napon. Ali u praksi se ne dostižu ni približno te vrednosti. Radi povećanja ključnih energetskih parametara, neophodno je obezbediti veću površinu elektroda, jer je glavni problem – relativno slab jonski transport koji još opada u toku svakog radnog punjenja baterije. Tako, na žalost, veća površina ne dovodi do očekivanih rezultata. Naime, prvenstveno – veoma je važno kako napraviti mikrostrukturu kompozita elektrode dostupnom za Li-jone, a da oni ne budu zarobljavani.

U tom smislu, nanostrukturni materijali u Li-jonskim baterijama igraju važnu ulogu. Naime, nanostrukturni materijali, zbog njihove posebne morfologije, pokazuju neočekivano elektrohemijsko ponašanje. Važno je napomenuti da, kada se nanostrukturni materijali koriste u Li-jonskim baterijama, njihove male dimenzije (do 10-tak nm) samo dovode do kraće dužine difuzije i veće površine Li inkorporiranja u čvrstu matricu. Ovaj problem se može rešiti upotrebom aktivnih materijala sa debljinom zidova od oko 10 – 20 nm [2]. Ovi materijali pozitivno utiču na jonski kontakt između aktivnih materijala i elektrolita. Jednaku zaslugu oni imaju i u postizanju boljeg elektronskog provođenja.

Dakle, za efikasno funkcionisanje Li-jon baterije potrebno je nesmetano paraleno provođenje elektronima i jonima, a za baterije velike snage ovakvo provođenje je i suštinsko. Ovde će biti analizirano nekoliko vidova nano-arhitektura specifičnog dizajna, prvenstveno od MO materijala.

Jovan P. Šetrajčić, Igor Šetrajčić i Ana Šetrajčić-Tomić – *Akademija nauka i umjetnosti Republike Srpske, Banja Luka, Republika Srpska, BiH* (e-mail: jovan.setrajcic@gmail.com, seki_1976@yahoo.com, setrajcic.a@gmail.com).

Siniša M. Vučenović – *Univerzitet u Banjoj Luci, Prirodno-matematički fakultet, Dr Mladema Stojanovića 2, 78000 Banja Luka, Republika Srpska, BiH* (e-mail: vucenovic.s@gmail.com).

Stevo K. Jaćimovski, *Kriminalističko-policijski univerzitet, Cara Dušana 196, 11080 Zemun – Beograd, Srbija* (e-mail: jacimovskis@gmail.com)..

Dušan I. Ilić i Nikola R. Vojnović – *Univerzitet u Novom Sadu, Fakultet tehničkih nauka, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija* (e-mail: dusilic69@gmail.com).

II. ELEKTRODE U LI-JONSKIM BATERIJAMA

Sve funkcionalne tehnologije baterija imaju kinetičkih problema sa opadajućom jonskom difuzijom kako u čvrstim Li elektrodama, tako i sa provodljivošću kroz elektrolite. U te svrhe, za poboljšanje jonske kinetike obično se predlažu specijalne arhitekture nano-strukturnih elektroda [2], koje se izrađuju u obliku dodatnog nanošenja ultratankih filmova na korišćene elektrode.

U toku rada baterije kontakt između elektroda i tih dodatih filmova treba da obezbedi mehaničku koheziju, ali i da pozitivno utiče na električne osobine elektroda preko površinske modifikacije. Korišćenjem mikroskopije atomskom silom (AFM), u radu [3] proučavane su površinski strukturni nanosi sa tri veziva: polivinildenfluorid (PVDF), carboksimethil celuloza (CMC) i želatin. Učinjeni su veliki napori da se pronađe veza između svojstava i arhitekture posmatranih struktura i elektrohemijskih karakteristika procesa punjenja–pražnjenja baterije.

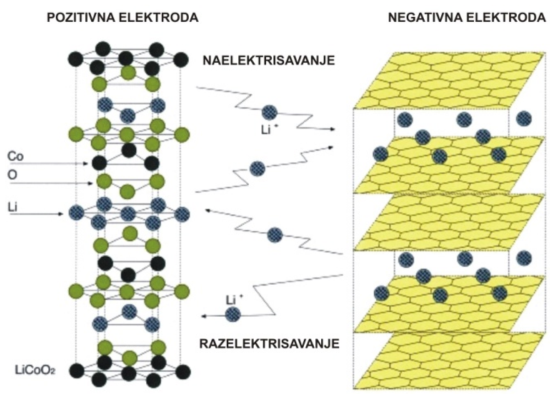
Poboljšanje jonske/elektronske provodnosti može se postići primenom više metoda, uključujući i nanošenje ugljenika, jonski super-valentni doping umesto Li i nano-umrežavanje elektronski provodnih metala. Pomoću ovoga može se dobiti dopirani materijal sa značajnom elektronskom provodljivosti – i do $4{,}8\cdot 10^{-2}$ S/cm. Iz ovih rezultata sledi da će i ostale osobine upotrebljenih materijala, kao što su fazne transformacije, imati značaan uticaj na ukupnu efikasnost [4].

$Li_{1-x}CoO_2$ je komercijalno najdostupniji materijal katode. Na žalost, njegova praktična primena je ograničena relativno velikom nestabilnošću, jer se brzo raspada već pri naponima iznad 4,2 V. Međutim, i ove nestabilnosti mogu se razrešiti upotrebom premaza $LiCoO_2$ praha nanošenjem MO nanoskopske debljine – do nekoliko desetina nm [5]. Primeri MO koji su istraživani su, npr. Al_2O_3, ZrO_2, ZnO, SiO_2, TiO_2,

$AlPO_4$ i AlF_3. Većina metoda formiranja prevlaka zasnivaju se na tehnikama kao što je *sol-gel*, ali se sve češće koristi metoda taloženja atomskog sloja (ALD – *Atomic Layer Deposition* [5]), koja je zapravo gasno–fazno formiranje ultratankih filmova.

III. ULOGA ULTRATANKIH PREMAZA ELEKTRODA

U najopštijem slučaju, efikasnost elektroda se može dovesti u direktnu vezu sa migracijom jona: što je ona veća, veća je i efikasnost elektrode (Sl. 1).



Sl. 1. Prikaz migracije jona u Li-jonskim baterijama (iz [3])
Fig. 1. Demonstration of ion migration in Li-ion batteries (from [3])

Opravdano se postavlja sledeće pitanje: na osnovu čega je povećana efikasnost elektroda posledica prisustva premaza, jer premaz, suštinski predstavlja dodatnu „barijeru" prolasku jona. Pokazaćemo da odgovor leži u izmenjenim mehaničim svojstvima premaza zbog delovanja „konfajnment" uslova. Naime, nije nemoguće napraviti totalnu analogiju tankih elektrodnih premaza i ultra-tankih MO filmova. Nakon uspostavljanja ove analogije, ostaje da se istraži uloga fononskog podsistema ultratankih filmova, a to je i učinjeno nadalje u ovom radu, jer je upravo taj sistem odgovoran za neobične mehaničke osobine.

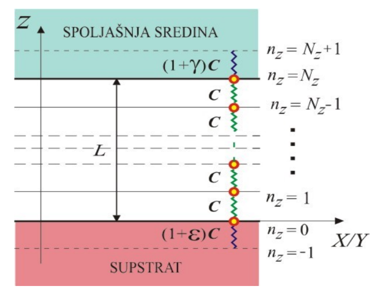
Modelni ultratanki filmovi [6–9] predstavljaju strukture sa narušenom translacionom simetrijom duž jednog pravca. Ovakve kristalne nanostrukture mogu se teorijski analizirati pomoću metoda sa jednočestičnim talasnim funkcijama, ali ovakav prilaz nije „zatvoren u sebe", jer se za izračunavanje statitičkih srednjih vrijednosti moraju „uzajmljivaiti" statističke formule. Jedini potpuno kompletan je metod Grinovih funkcija (GF) [10–15], jer on pruža mogućnost za ozračunavanje dinamičkih i statističkih karakteristika posmatranog sistema. Kao prvo i osnovno, ovaj metod smo uspešno „prilagodili" za istraživanje nanoskopskih kristalnih struktura [16–20] u kojima egzistiraju nanoskopske dimenzije i postoje različite granične površi, poznate kao „konfajnment" uslovi, koji unose najbitnije promene u sve fundamentalne osobine date strukture.

IV. MEHANIČKO OSCILOVANJE U FILM-STRUKTURI

Hamiltonijan mehaničkih oscilacija u ultratankom filmu (Sl. 2) uzet u aproksimaciji najbližih suseda i uz zanemarivanje torzionih efekata je oblika

$$H = \frac{1}{2M}\sum_{\vec{n}} p_{\vec{n}}^2 + \frac{C}{2}\sum_{\vec{n}}\Big[\left(u_{\vec{n}} - u_{\vec{n}-\vec{a}_x}\right)^2 + \left(u_{\vec{n}} - u_{\vec{n}-\vec{a}_y}\right)^2 + \left(u_{\vec{n}} - u_{\vec{n}-\vec{a}_z}\right)^2\Big], \qquad (1)$$

gde su u – pomjeraji, p – impulsi, M – mase molekula i C – Hukove konstante istezanja. Torzioni efekti mogu biti isključeni, jer se fundamentalni efekat postiže u pravcu normalnom na granične površi strukture, dakle duž pravca gde je posmatrana struktura debljine do desetak nm i zanemariva je u odnosu na druge dve dimenzije!



Sl. 2. Model fononskog ultratankog filma (iz [15])
Fig. 2. Model of phonon ultrathin film (from [15])

Posmatraćemo GF tipa pomeraj-pomeraj, koja se definiše [13–15], na sedeći način:

$$\Psi_{\vec{n},\vec{m}}(t) \equiv \Psi_{n_x,n_y,n_z;m_x,m_y,m_z}(t) \equiv \langle\langle u_{\vec{n}}(t) | u_{\vec{m}}(0) \rangle\rangle = \theta(t) \langle [u_{\vec{n}}(t), u_{\vec{m}}(0)] \rangle, \qquad (2)$$

a GF tipa impuls-impuls:

$$\Phi_{\vec{n},\vec{m}}(t) \equiv \Phi_{n_x,n_y,n_z;m_x,m_y,m_z}(t) \equiv \langle\langle p_{\vec{n}}(t) | p_{\vec{m}}(0) \rangle\rangle = \theta(t) \langle [p_{\vec{n}}(t), p_{\vec{m}}(0)] \rangle. \qquad (3)$$

Pri rešavanju sistema jednačina za ove GF iskorišćena je činjenica da je film dimenziono neograničen u x i y pravcu i da je translatorno invarijantan. Međutim, duž pravca narušenja simetrije (z) granični uslovi postoje i moraju se pravilno definisati:

$$u(n_z = -1) = u(n_z = N_z - 1) = 0; \quad p(n_z = -1) = p(n_z = N_z - 1) = 0. \qquad (4)$$

Zbog ovakvih graničnih uslova jednačine za određivanje Ψ i Φ „raspadaju se" na sistem od tri jednačine [13]. Zbog toga, komponente funkcija Ψ i Φ zavise od indeksa sloja n_z, pa se moraju razvijati po stojećim talasima oblika $\sin(N_z + 2)\varphi_\mu$, gde je parametar φ_μ definisan kao:

$$\varphi_\mu = \frac{\pi\mu}{N_z + 2}; \quad \mu = 1,2,\ldots,N_z + 1. \qquad (5)$$

Veoma je bitna činjenica ta da, u formuli (5), broj μ ne može da uzme vrijednost 0, a ni N_z+2, jer bi tada GF Ψ i Φ bile jednake nuli, a to je neinteresantan – trivijalni slučaj.

Opisanim postupkom određene su GF Ψ i Φ, a njihovi polovi u ω ravni, koji određuju energetsa stanja elementarnih pobuđenja – ovde fonona, definisani su sledećom kvadratnom determinantom:

$$D_{N_z+1}(\rho) = \begin{vmatrix} \rho-\varepsilon & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & \rho & 1 & \dots & 0 & 0 & 0 \\ 0 & 1 & \rho & \dots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \ddots & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & \rho & 1 & 0 \\ 0 & 0 & 0 & \dots & 1 & \rho & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & \rho-\gamma \end{vmatrix}_{N_z+1}$$

gde je

$$\rho_k^\alpha = \frac{\omega^2}{\Omega_\alpha^2} - 4\sin^2\frac{ak_x}{2} - 4\sin^2\frac{ak_y}{2} - 2 \equiv \rho,$$

Ali i sledećim uslovom:

$$D_{N_z+1}(\rho) = 0. \tag{6}$$

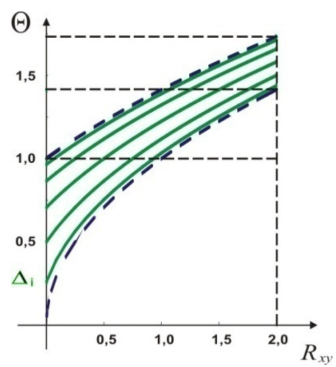
U sasvim opštem slučaju, uslov (6) nije egzaktno rešiv, jer njegov red, kao i rešenje zavise od tri parametra: ε, γ i N_z, čeje vrednosti diktira sama modelna struktura, njena arhitektura, ali i precifična kontaktna svojstva supstrata, ovde elektroda i lepka, i okoline, ovde elektrolita.

Egzaktno rešenje ipak postoji, u potpuno idealizovanom slučaju [16], kada parametri filma imaju vrednost: $\varepsilon = \gamma = 0$. Tada se zakon disperzije (spektar dozvoljenih energija) fonona u filmu dobija u obliku:

$$\Theta \equiv \frac{E}{\hbar\Omega} = 2\sqrt{\sin^2\frac{a_x k_x}{2} + \sin^2\frac{a_y k_y}{2} + \sin^2\frac{\pi\mu}{2(N_z+2)}}; \tag{7}$$

$$\Omega = \sqrt{\frac{C}{M}}; \quad \mu = 1,2,\dots,N_z+1.$$

Ovaj zakon disperzije grafički je predstavljen na Sl. 3, gde je: $R_{xy} = \sin^2\frac{a_x k_x}{2} + \sin^2\frac{a_y k_y}{2}$. Uočljive su diskretne vrednosti mogućih energija.



Sl. 3. Zakon disperzije fonona u „idealnom“ filmu
Fig.. 3. Dispersion law of phonons in „ideal“ film

Iako je ovo idealizovan slučaj i udaljen od prakse, on se mora izučiti, iz više razloga. Prvi je taj koji će nam omogućiti već nekakvo zaključivanje, a to je odsustvo kontinualnosti, tj. postojanje izrazite diskretnosti i prebrojive brojnosti energetskih fononskih stanja. Drugi razlog proizilazi iz činjenice – rešenja za GF, da se između dve granične površi ne prostire običan mehanički (zvučni) talas, nego specifičan stojeći, koji ima dvostruko veću amplitudu! Treći razlog je da u proračunima koji slede i koje ćemo morati numerički da rešavamo, imamo „kontrolnu tačku“, kada drastično smanjimo vrednosti graničnih parametara. Dakle, svi ostali – neidealni slučajevi urađeni su numerički, upotrebom programskog paketa *Mathematica*. Grafici su slični onom prikazanom na pređašnjoj Sl. 3, ali se raspored (zelenih) fononskih stanja menja u zavisnosti od veličine graničnih parametara.

Evidentno je i to da u ovakvom sistemu ne mogu da nastanu izolovana, tj. lokalizovana fononska stanja, njihov broj ostaje uvek isti i unutar „balkovskih“ granica, a definisan je brojem kristalnih ravni duž pravca narušenja translacione simetrije.

Nakon ovoga, uobičajenom procedurom [13] određene su srednje vrednosti kvadrata pomeraja i kvadrata impulsa i one su bile date forulama:

$$< u_{\vec n}^2 >= \frac{\hbar}{M}\frac{1}{N_x N_y}\frac{1}{N_z+2}\sum_{k_x,k_y}\sum_{\mu=1}^{N_z+1}\frac{1}{\omega_{k_x,k_y,\mu}}\cdot \tag{8}$$

$$\cdot \sin^2\left(n_z+1\right)\frac{\pi\mu}{N_z+2}\coth\frac{\hbar\omega_{k_x,k_y,\mu}}{2\Theta};$$

$$< p_{\vec n}^2 >= \frac{\hbar M}{N_x N_y\left(N_z+2\right)}\sum_{k_x,k_y}\sum_{\mu=1}^{N_z+1}\omega_{k_x,k_y,\mu}\cdot \tag{9}$$

$$\cdot \sin^2\left(n_z+1\right)\frac{\pi\mu}{N_z+2}\coth\frac{\hbar\omega_{k_x,k_y,\mu}}{2\Theta}.$$

Iz dobijenih izraza se vidi da, za razliku od idealne strukture, srednji kadrati pomeraja i impulsa zavise od prostorne koordinate n_z.

Moramo podvući, da je za nas veoma bitno i evidentno da rešenja GF uvek predstavljaju stojeće talase, slično kao kod vazdušnog stuba. Pri tome, trbusi se nalaze na graničnim ravnima filma, što ukazuje da atomi na tim ravnima imaju najveću energiju i amplitudu oscilovanja.

V. Zaključna analiza rezultata

U fononskom podsistemu ultratankih kristalnih filmova, koji je odgovoran, odnosno – koji opisuje mehanička svojstva posmatrane strukture, dolazi do pojave znatno pojačanog oscilovanja i formiranja stojećih talasa. Pojačanim fononskim delovanjem – za očekivati je, i njihovo vrlo intenzivno i pozitivno delovanje na veličinu jon-fonon interakcije, jer najveći deo njihove – mehaničke energije, ide upravo na ovo povećanje. Usled toga može se jasno zaključiti da bi odgovor na pitanje mehanizma za poboljšanje efikasnosti jonskog transporta, upravo mogao da bude u značajnom pozitivnom uticaju vibracija kristalne rešetke prevlaka elektroda na „razbijanje“ klopki koje zarobljavaju jone u graničnom sloju elektroda. Ove znatno jače vibracije – na neki poseban način „razmrdavaju” ili „oživljavaju” jone, i na taj način im stvaraju uslove za brojniji, neometaniji i, u svakom slučaju, mnogo brži transfer kroz elektrolit, sa jedne na drugu elektrodu i nazad.

Istovremeno, svojim znatno ojačanim oscilovanjem, fononi oslobađaju u elektrodama i u zaprečnom međusloju između graničnog sloja elektroda i elektrolita, one „zarobljene” jone koji, bez prisustva predloženih i ovde analiziranih MO prevlaka, ne bi više nikako mogli da učestvuju u prenosu naelektrisanja i energije između elektroda unutar elektrolita u Li-jonskim baterijama. Na taj način, sistem: elektrode sa ultratankim MO filmovima – u mogućnosti je da značajnije i trajnije povećava efikasnosti Li-jon provođenja.

ZAHVALNICA

Istraživanja čiji su rezultati ovde prezentovani, finanskijski su potpomognuta od Ministarstva za naučnotehnološki razvoj, visoko obrazovanje i informaciono društvo Republike Srpske (Projekti br. 19/6-020/961-2/18, 19/6-020/961-35/18, 19.032/961-36/19 i 19.032/961-42/19).

LITERATURA

[1] J.-M. Tarascon, M. Armand, Nature, 414, (2001) 359-367

[2] P. L. Taberna, S. Mitra, P. Poizot, P. Simon & J.-M. Tarascon, Nature Materials 5, 567 – 573 (2006).

[3] S. Pejovnik, R. Dominko , M. Bele, M. Gaberscek and J. Jamnik, Journal of Power Sources, Vol. 184, Issue 2, 593-597 (2008).

[4] J. Hong, Ch. Wang , U. Kasavajjula, Journal of Power Sources 162 (2006) 1289–1296.

[5] Yoon Seok Jung, Andrew S. Cavanagh,b Anne C. Dillon, Markus D. Groner, Steven M. George, and Se-Hee Leea, Journal of The Electrochemical Society*,* 157, 1, A75-A81, 2010.

[6] D. Popov, S.K. Jaćimovski, B.S. Tošić, J.P. Šetrajčić, Physica A vol.317, 129-139 (2003).

[7] B.S. Tošić, J.P.Šetrajčić, D.Lj. Mirjanić, Z.V. Bundalo, Physica A vol.184, 354-366 (1992).

[8] S.G. Davison and M. Steslicka: Basic Theory of Surface States, Clarendon , Osford (1996).

[9] M. Prutton: Introduction to Surface Physics, Clarendon, Oxfod (1995).

[10] M. Tkach, V. Holovatsky, O. Voitsekhivska, Physica E vol.11, 17-26 (2001).

[11] V.M. Golovach, G.G. Zegrya, A.M. Makhanets, I.V. Pronishin, N.V. Tkach, Semiconductors, vol.33/5, 564-568 (1999).

[12] J.M.Wesselinowa, Phys.Stat.Sol. (b), vol.223, 737 (2001); vol.229, 1329 (2002).

[13] J.P.Šetrajčić, S.K.Jaćimovski, Review of Results of Theoretical Approaches to Phonon Engineering of Thermodynamic Properties for Different Quantum Structures, *NBP* **20**/3, 67-82 (2015).

[14] J.P.Šetrajčić, S.K.Jaćimovski, D.I.Ilić, Process of Phonon Diffusion through Crystalline Structures, *NBP* **22**/2, 1-12 (2017).

[15] B.S.Tošić, J.P.Šetrajčić, S.K.Jaćimovski: Metodi teorijske fizike, Kriminalističko-policijski univerzitet, Zemun 2018.

[16] M. Mozetič et all: Recent developments in surface science and engineering, thin films, nanoscience, biomaterials, plasma science, and vacuum technology, Thin Solid Films **660**, 120-160 (2018)

[17] B.S. Tošić*:* Statisticka fizika, PMF, Novi Sad (1978).

[18] S. Doniach, E.H. Sondheimer: Green's Functions for Solid State Physicists, Imperial College Press, London (1999).

[19] D.Lj.Mirjanić and J.P.Šetrajčić, Proceedings UNITECH'08 **3**, 566-571 (2008).

[20] C. Kittel, Introduction to Solid State Physics, Wiley, New York (1986).

ABSTRACT

The paper presents the results of theoretical research of the phonon subsystem behavior in ultrathin metal-oxide (MO) coatings when they are applied to electrodes in Li-ion batteries in order to increase the efficiency of ion transport. The method of Green's functions was applied and numerical analysis showed that in ultra-thin films (in direction perpendicular to the boundary surfaces of the film) occur distinctly amplified mechanical oscillation of the crystal lattice and the formation of standing waves, whereby phonons stimulate the release of trapped ions in the boundary layers of the electrodes and with that increase the efficiency of ion transport between the electrodes, and thus contribute to increasing of the electrical conductivity in these batteries.

Increase in the Electrical Conductivity of Li-ions Batteries by Electrode Coating Metal-Oxide Nano-Films

Jovan P. Šetrajčić, Siniša M. Vučenović,
Igor J. Šetrajčić, Stevo K. Jaćimovski,
Ana J. Šetrajčić-Tomić, Dušan I. Ilić, Nikola R. Vojnović

Performanse sklopova termoelektrični modul - hladnjak namenjenih samonapajajućim sistemima u uslovima prirodnog hlađenja

Aleksandra Stojković, *Student Member, IEEE*, Miloš Marjanović, *Student Member, IEEE*, Jana Vračar, *Student Member, IEEE*, Aneta Prijić, *Member, IEEE* i Zoran Prijić, *Member, IEEE*

***Apstrakt*— Sklopovi termoelektrični modul-hladnjak nalaze široku primenu u sistemima za konverziju termičke u toplotnu energiju i obrnuto. U ovom radu je analiziran izabrani termoelektrični modul u ulozi termoelektričnog generatora u sprezi sa šest različitih hladnjaka bliskih dimenzija pod uslovima prirodnog hlađenja. Razmatrana je efikasnost sklopova sa aspekta električnog napona predatog potrošaču pri različitim temperaturnim pobudama. Izvršena su eksperimantalna merenja i odgovarajuće numeričke simulacije, pri čemu su analizirani doprinosi pojedinih termoelektričnih efekata i mehanizama odvođenja toplote sa površina hladnjaka. Pokazano je da ravni hladnjaci od mikroporoznih materijala mogu adekvatno da zamene ekstrudirane aluminijumske hladnjake kod samonapajajućih sistema u uslovima prirodnog hlađenja.**

***Ključne reči*—Termoelektrični generator; hladnjak; samonapajući sistem; prirodno hlađenje; radijacija.**

I. UVOD

Termoelektrični moduli (TEM-ovi) u sprezi sa različitim tipovima hladnjaka se često susreću kao delovi prenosnih mini uređaja za hlađenje/grejanje ili termoelektrično napajanih bežičnih senzorskih čvorova. Sami moduli baziraju svoj rad na Peltijevom (Peltier) odnosno Zebekovom (Seebeck) efektu. Kada rade kao izvori ili apsorberi toplote termoelektrični moduli se pobuđuju električnom strujom i kao takvi su poznati pod nazivom Peltijevi elementi za grejanje odnosno hlađenje [1]. Analogno, kada se na suprotne strane ovih modula primeni odgovarajuća temperaturna razlika oni generišu električni napon i tada se klasifikuju kao termoelektrični generatori (TEG-ovi) [2]. Za efikasan rad TEM-ova je neophodna dopunska razmena toplote sa njegove hladne i/ili tople strane što se obezbeđuje kako aktivnim tako i pasivnim hladnjacima. S obzirom da su standardni pasivni hladnjaci znatno jednostavnije konstrukcije i bez potrebe za dodatnim podsklopovima za protok rashladnog fluida, oni su široko prisutni u komercijalnim sistemima. Performanse sklopova sa pasivnim hladnjacima, pod definisanim radnim uslovima, osim od karakteristika TEM-a zavise i od konstrukcije i materijala upotrebljenog hladnjaka [3].

Aluminijumski hladnjaci sa ekstrudiranim rebrima ili stubićima su standardno rešenje za primene u kojima se zahteva jednostavno povećanje površine za odvođenje toplote. Hladnjaci od keramike alumine (Al_2O_3) zamenjuju aluminijumske pri zahtevima za visokim probojnim naponima i elektromagnetnom kompatibilnošću i primenjuju se kod LED rasvete, modula snage i specifičnih integrisanih kola [4]. Mikroporozni keramički hladnjaci ravne ili jednostavne rebraste geometrije su niskoprofilni, lagani i bez efekta antene [5]. Njihova mikroporozna struktura obezbeđuje povećanje efektivne kontaktne površine sa fluidom i niski toplotni kapacitet po jedinici zapremine, što ih čini pogodnim za primenu u sistemima gde prevladava prirodno hlađenje (bez strujanja okolnog fluida) unutar ograničenog prostora. Odnedavno komercijalno dostupne hladnjake od bakarne pene takođe karakteriše mikroporozna struktura, kao i dodatna prevlaka od bakar-oksida radi povećanja emisivnosti [6].

Ovaj rad analizira funkcionisanje sklopova TEG-hladnjak prvenstveno za primenu u bežičnim senzorskim čvorovima u uslovima prirodnog hlađenja. Osnovni cilj je da se uporede performanse nestandardnih pasivnih hladnjaka (keramičkih i mikroporoznih) sa aluminijumskim hladnjacima bliskih spoljašnjih dimenzija. Osim vrednosti generisanog napona na potrošaču, prati se i vremenski odziv sklopa na iznenadnu termičku pobudu, odnosno njegova termička inertnost. Sva eksperimentalna merenja prate numeričke simulacije sa ciljem uspostavljanja standardizovane procedure za predviđanje karakteristika novoprojektovanih sklopova.

II. EKSPERIMENTALNA POSTAVKA

Razmatraju se komercijalno dostupne komponente tako da je izabran tipičan „minijaturni“ TEG GM200-127-14-16 sa 127 termoelektričnih parova, spoljašnjih dimenzija (40×40×3,8) mm [4]. Karakteristični električni parametri ovog TEG-a su dati u Tabeli I. Na hladnu stranu TEG-a je termoprovodnom adhezivnom trakom pričvršćen hladnjak. Analizirno je šest niskoprofilnih hladnjaka čije su osnove dimenziono kompatibilne sa pločama TEG-a a izrađeni su od aluminijuma, alumine, mikroporozne keramike i bakarne pene. Parametri izabranih hladnjaka su navedeni u Tabeli II.

Aleksandra Stojković – Univerzitet u Nišu, Elektronski fakultet, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: aleksandra.stojkovic@elfak.ni.ac.rs).

Miloš Marjanović – Univerzitet u Nišu, Elektronski fakultet, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: milos.marjanovic@elfak.ni.ac.rs)

Jana Vračar – Univerzitet u Nišu, Elektronski fakultet, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail:jana.vracar@ elfak.ni.ac.rs)

Aneta Prijić – Univerzitet u Nišu, Elektronski fakultet, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: aneta.prijic@elfak.ni.ac.rs)

Zoran Prijić – Univerzitet u Nišu, Elektronski fakultet, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: zoran.prijic@elfak.ni.ac.rs)

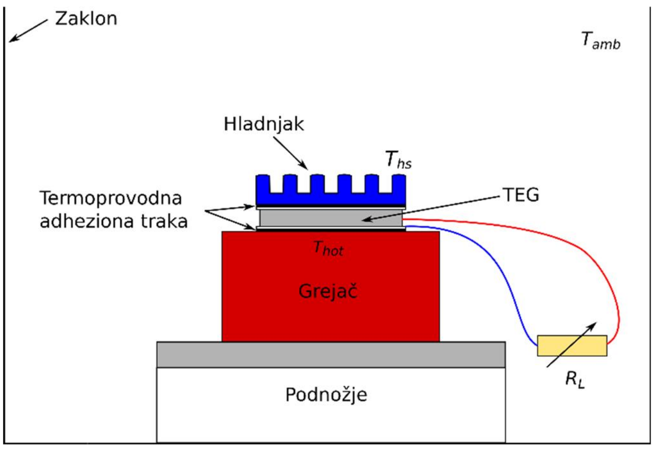
Eksperimentalna postavka koja obezbeđuje automatizovanu kontrolu radnih uslova i merenje napona predatog od strane TEG-a potrošaču je šematski prikazana na Sl. 1. Sklop je termoprovodnom trakom fiksiran za površinu grejača u horizontalnom položaju a ceo sistem je okružen zaklonom koji obezbeđuje uslove prirodnog hlađenja. Grejač generiše toplotnu pobudu, pri čemu se njegova temperatura, koja je ujedno i temperatura tople strane TEG-a (T_{hot}), zadaje po predefinisanim vrednostima nekoliko desetina °C višim od temperature okoline (T_{amb}). S obzirom da nije moguće direktno merenje temperature hladne strane TEG-a (T_{cold}), kao kontrolni parametar se prati temperatura površine hladnjaka (T_{hs}). Električni izvodi TEG-a su vezani za programabilno aktivno opterećenje koje istovremeno predstavlja potrošač promenljive vrednosti (R_L) i ima ulogu multimetra. Sve vrednosti temperatura se prate pomoću računarom vođenih digitalnih multimetara.

TABELA I
KARAKTERISTIČNI ELEKTRIČNI PARAMETRI TEG-A GM200-127-14-16

Maksimalna snaga - P_{max}	4,73 W
Maksimalna struja - I_{max}	1,19 A
Maksimalni napon – U_{max}	3,98 V
Unutrašnja otpornost - R_{TEG}	2,1 Ω

TABELA II
GEOMETRIJSKI PARAMETRI RAZMATRANIH HLADNJAKA

Oznaka hladnjaka	Geometrija	Materijal	Dimenzije L×W×H mm
HS1	110 stubića	aluminijum	40×40×10
HS2	100 stubića	aluminijum	40×40×5
HS3	121 stubić	alumina	42×42×12
HS4	7 niskih rebara	mikroporozna keramika	40×40×5,25
HS5	ravan	mikroporozna keramika	40×40×2,5
HS6	ravan	bakarna pena	40×40×5



Sl. 1. Šematski prikaz eksperimentalne postavke za kontrolu radnih uslova sklopa i merenje temperature i napona na potrošaču.

Grejač se postavlja na željenu temperaturu i prati promena T_{hot} kao i napon otvorenog kola TEG-a (V_{oc}). Kada sistem uđe u termičku ravnotežu, uključuje se aktivno opterećenje koje obezbeđuje otpornost u opsegu od 2 Ω do 5 Ω sa korakom od 1 Ω (od veće ka manjoj vrednosti). Ove otpornosti odgovaraju ulaznoj otpornosti tipičnog kola za upravljanje energijom u samonapajajućim sistemima. Za svaku zasebnu vrednost otpornosti potrošača meri se vrednost pada napona na njemu (V_L) i struje kroz kolo (I_L). Zatim se isključuje aktivno opterećenje i kada se napon otvorenog kola vrati na stacionarnu vrednost ponavlja se ciklus za sledeću temperaturu grejača. S obzirom da je opterećenje direktno vezano za TEG, električne veličine zadovoljavaju relaciju:

$$\begin{aligned} V_L &= R_L I_L = V_{TEG} - R_{TEG} I_L = \\ &= N\alpha_{pn}(T_{hot} - T_{cold}) - R_{TEG} I_L, \end{aligned} \tag{1}$$

gde je N broj termoelektričnih parova a α_{pn} Zebekov koeficijent materijala TEG-a.

III. POSTAVKA SIMULACIJE

Postavka simulacije u potpunosti prati eksperimentalnu postavku sa Sl. 1. Elementi sklopa i okruženje su modelirani u CAD softveru [7] i uvezeni u simulator za numeričku multifizičku analizu metodom konačnih elemenata [8]. Simulira se stacionarno stanje pri odgovarajućoj temperaturnoj pobudi primenom modula za rešavanje problema u spregnutom termo-električnom i domenu dinamike fluida. Simulacioni domen obuhvata okolni fluid (vazduh) dovoljnih dimenzija da granični uslovi odgovaraju uslovima prirodnog hlađenja. Sa gornje strane je domen otvoren kako bi se dozvolilo slobodno odvođenje toplote.

Model TEG-a obuhvata 127 termoelektričnih spojeva (parova Bi_2O_3 nožica n- i p-tipa) redno povezanih bakarnim kontaktima i smeštenih između dve ploče od alumina keramike. Geometrija hladnjaka je modelirana detaljno, dok su grejač i podnožje pojednostavljeni. Termoprovodna adheziona traka se, usled male debljine, smatra idealnim termičkim kontaktom. Grejač obezbeđuje konstantan dotok toplote kao uslov termičkog opterećenja dok eksterni otpornik vezan za izvode TEG-a predstavlja električno opterećenje. Uključeni mehanizmi prenosa toplote su provođenje kroz čvrste i mikroporozne strukture, laminarno strujanje zagrejanog fluida pod dejstvom gravitacije i radijacija. U uslovima prirodnog hlađenja, radijacija ima značajan doprinos odvođenju toplote sa elemenata sklopa kako putem direktnog zračenja površine ka ambijentu tako i putem indirektnog zračenja od jedne ka drugoj površini [9]. Automatski kreirana inicijalna simulaciona mreža je naknadno poboljšana u okolini dodirne površine hladnjaka sa vazduhom, tako da tipična mreža sadrži 400.000-500.000 diskretizacionih elemenata.

Specifični geometrijski i parametri materijala TEG-a i hladnjaka, neophodni za realizaciju simulacije, dati su u Tabeli III i Tabeli IV, respektivno. Naglašava se da su ukupan Zebekov koeficijent, specifična električna otpornost i specifična termička provodnost Bi_2Te_3 nožica temperaturno zavisni parametri [10]. Kod mikroporoznih hladnjaka, usled složene strukture uslovljene postojanjem otvorenih pora, za proračun efektivnih vrednosti specifične termičke provodnosti, specifičnog toplotnog kapaciteta, kao i efikasnosti protoka fluida, softver koristi vrednosti koeficijenta poroznosti i permeabilnosti [11].

TABELA III
GEOMETRIJSKI I PARAMETRI MATERIJALA TEG-A

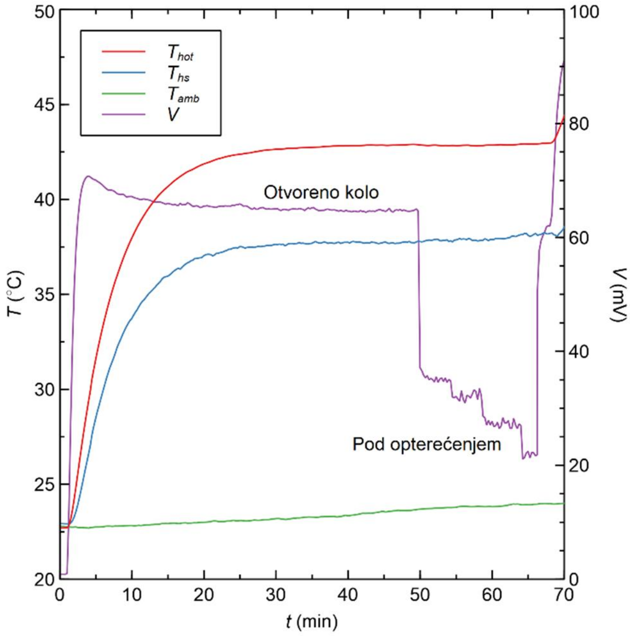
Dimenzije nožica - l×w×h (mm)	1,4×1,4×1,6
Dimenzije Cu kontakta (mm)	4,2×1,4×0,3
Debljina ploče od alumine (mm)	0,8
Ukupan Zebekov koeficijent termopara α_{pn} (μV/K)	$1{,}22\cdot10^{-5}\times T^3$-$0{,}021\times T^2$+$10{,}23\times T$-1081
Specifična električna otpornost termopara (10^{-5} Ωm)	$-1{,}48\cdot10^{-5}\times T^2$+$0{,}017\times T$-2,78
Specifična termička provodnost termopara (W/mK)	$3{,}96\cdot10^{-5}\times T^2$-$0{,}026\times T$+5,84

TABELA IV
PARAMETRI MATERIJALA RAZMATRANIH HLADNJAKA

Parametar	Materijal			
	Aluminijum	Alumina	Mikro-porozna keramika	Bakarna pena
Gustina (kg/m^3)	2700	3660	1800	3320
Specifična termička provodnost (W/mK)	201	25	125	45
Specifični toplotni kapacitet (J/kgK)	900	880	670	385
Poroznost (%)	-	-	30	63
Permeabilnost (m^2)	-	-	$0{,}54\cdot10^{-10}$	$5{,}32\cdot10^{-8}$
Emisivnost	0,85	0,6	0,6	0,7

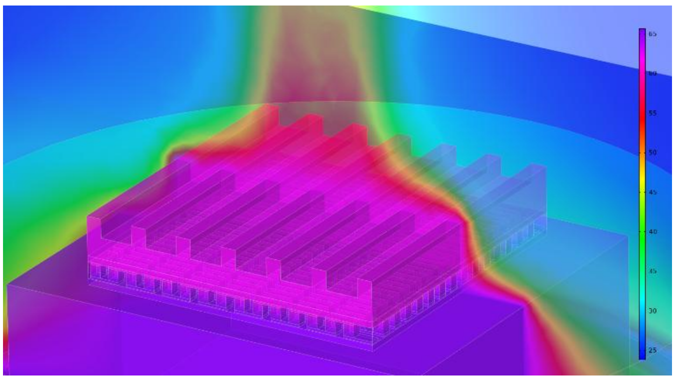
IV. REZULTATI I DISKUSIJA

Tipičan rezultat eksperimentalnog merenja na sklopu sa hladnjakom HS4 je prikazan na Sl. 2. Na sklop se primenjuje iznenadna termička pobuda ΔT=T_{hot}-T_{amb}=20 °C koja stvara temperaturnu razliku na stranama TEG-a a time se generiše napon otvorenog kola V_{oc}=V_{TEG}. Ovakva pobuda odgovara uslovima hladnog starta u samonapajajućim sistemima. U prvom trenutku postoji premašenje napona koje nastaje usled termičke inertnosti sklopa [12] i ono lagano opada kako se sklop zagreva i pune njegove termičke kapacitivnosti [13]. Kada sklop uđe u stabilno stanje uključuje se aktivno opterećenje tako da kroz TEG protiče struja I_L a vrednost merenog napona V_L prati relaciju (1). Treba napomenuti da je napon V_{TEG} pri uslovima proticanja struje manji od V_{oc} usled izraženog delovanja Peltijevog efekta koji povećava temperaturu hladne a smanjuje temperaturu tople strane, odnosno smanjuje efektivnu temperaturnu razliku na TEG-u [10]. Smanjenje napona je izraženije pri većim temperaturnim razlikama ΔT i nižim vrednostima R_L usled viših vrednosti I_L. Ovaj fenomen se može kvalitativno uočiti pri isključenju aktivnog opterećenja kod sklopova sa aluminijumskim hladnjacima jer su oni manje inertni i lakše se oslobode toplote generisane Peltijevim efektom.



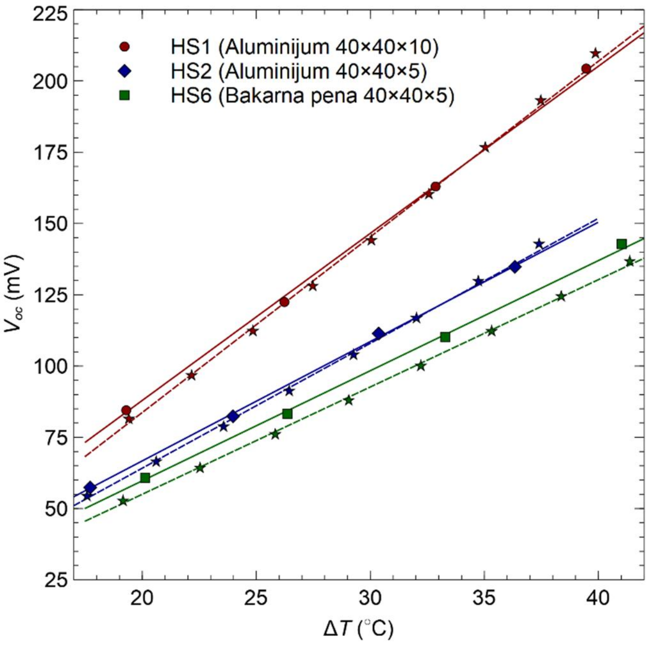
Sl. 2. Eksperimentalne vrednosti temperature na sklopu sa HS4 i napona na potrošaču za ΔT=T_{hot}-T_{amb}=20 °C.

Temperaturna raspodela dobijena simulacijom za sklop sa HS4 pri ΔT=T_{hot}-T_{amb}=40 °C i bez električnog opterećenja TEG-a je prikazana na Sl. 3. Jasno se uočava raspodela temperature kroz nožice i na hladnoj strani TEG-a kao i profil strujanja toplog vazduha sa hladnjaka, odnosno oblik perjanice. Simulacijama je uočeno da efekat radijacije, značajno doprinosi izgledu raspodele i vrednostima temperature na elementima sklopa.



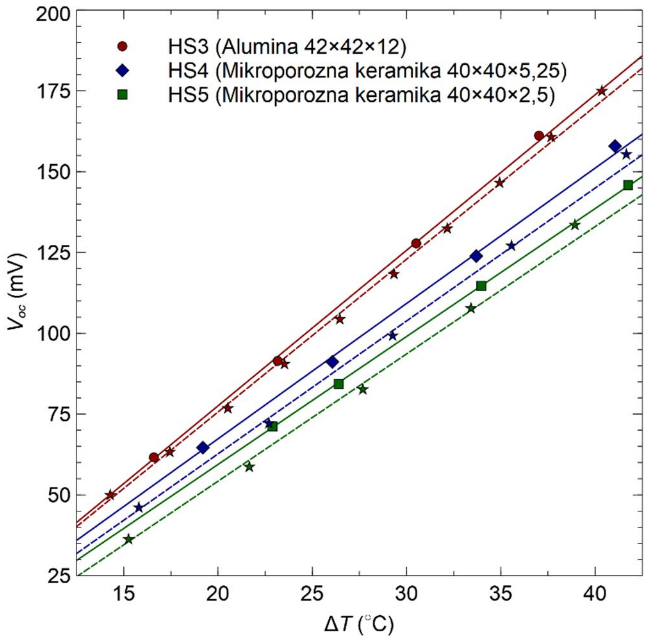
Sl. 3. Raspodela temperature na sklopu sa HS4 i okolnom fluidu dobijena simulacijom pri ΔT=T_{hot}-T_{amb}=40 °C i pri otvorenom kolu.

Rezultati eksperimentalnih merenja i simulacija su predstavljeni u vidu zavisnosti napona V_{oc} od temperaturne razlike $\Delta T{=}T_{hot}{-}T_{amb}$ na Sl. 4 i Sl. 5. Opseg temperaturnih razlika je karakterističan za najčešće dostupne razlike u temperaturi grejača i okoline kod samonapajajućih sistema. Rezultati simulacije veoma dobro prate rezultate eksperimentalnih merenja. Maksimalno odstupanje kod aluminijumskih i alumina hladnjaka je 6% dok je kod mikroporoznih hladnjaka 10%, prvenstveno zbog nemogućnosti softvera da adekvatno simulira efekte indirektne radijacije kod ovakvih materijala. Napominje se da, obično zanemarivana, radijacija utiče čak do 30% na vrednost generisanog napona dobijenu simulacijama.



Sl. 4. Zavisnost napona otvorenog kola od temperaturne razlike $\Delta T{=}T_{hot}{-}T_{amb}$ za sklopove sa hladnjacima od aluminijuma i bakarne pene. Simboli i pune linije – eksperiment; zvezdice i isprekidane linije - simulacije.

Hladnjaci HS1 i HS3 daju najviše vrednosti napona za ceo opseg razmatranih temperaturnih razlika, što je i očekivano s obzirom da imaju najveći broj i visinu stubića ali oni zauzimaju najveću zapreminu. Uočava se da hladnjaci HS2 i HS4, bliskih spoljašnjih gabarita, imaju skoro identičnu efikasnost. Ovo je značajno jer je HS2 ekstrudirani aluminijumski hladnjak, složene geometrije a HS4 od mikroporozne keramike, niskoprofilni i jednostavne geometrije. Mikroporozni ravni hladnjaci HS5 i HS6 su skoro jednakih performansi. Pri tome, debljina i poroznost hladnjaka od keramike je dvostruko manja u odnosu na hladnjak od bakarne pene. Kvantitativnim poređenjem karakteristika hladnjaka HS2 i HS6, uočava se da aluminijumski hladnjak daje približno 10% više napone, što ne predstavlja prevelik doprinos. Istovremeno, analiza rezultata za mikroporozne keramičke hladnjake HS4 i HS5, ukazuje da dvostruko veća debljina hladnjaka poboljšava performanse za svega 8%. Opšti zaključak je da u uslovima prirodnog hlađenja, a uzevši u obzir gabarite i složenost geometrije, keramički materijali (alumina i mikroporozna keramika) pokazuju najbolju efikasnost. Ravni hladnjaci od bakarne pene se takođe mogu uspešno primeniti u uslovima prirodnog hlađenja. Evidentno je da stubići kod ekstrudiranih hladnjaka poboljšavaju efikasnost prevashodno kod hlađenja uz postojanje strujanja okolnog fluida.

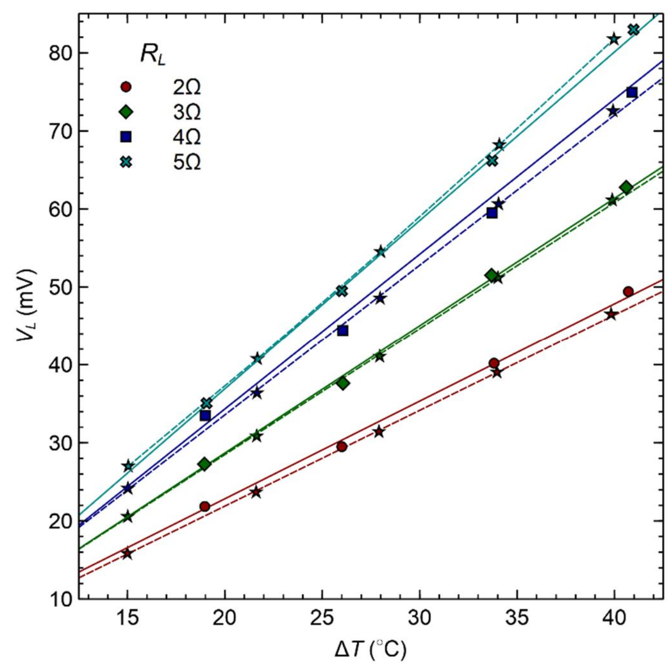


Sl. 5. Zavisnost napona otvorenog kola od temperaturne razlike $\Delta T{=}T_{hot}{-}T_{amb}$ za sklopove sa hladnjacima od alumine i mikroporozne keramike. Simboli i pune linije – eksperiment; zvezdice i isprekidane linije - simulacije.

Zavisnost napona na potrošaču V_L od temperaturne razlike za sklop sa rebrastim hladnjakom od mikroporozne keramike (HS4) i pri različitim vrednostima aktivnog opterećenja R_L je prikazana na Sl. 6. Rezultati omogućavaju kvantitativnu procenu uticaja Peltijevog efekta na vrednost generisanog napona. U odnosu na vrednost V_L koja bi se dobila raspodelom V_{oc} na otpornostima R_L i R_{TEG}, ovaj efekat doprinosi njenom smanjenju za 22% do 36% u zavisnosti od ΔT i R_L, što je u saglasnosti sa ranijim analizama termoelektrično napajanih senzorskih čvorova [14].

V. ZAKLJUČAK

Analiza performansi sklopova TEM-hladnjak treba da obezbedi jednostavniji proces projektovanja sistema koji se na njima baziraju. U ovom radu je akcenat stavljen na primene u termoelektrično napajanim bežičnim senzorskim čvorovima gde moduli imaju ulogu generatora i to pod najnepovoljnijim uslovima rada tj. pri prirodnom hlađenju. Posvećena je pažnja i radu sklopova pod opterećenjem koji odgovaraju realnom funkcionisanju senzorskog čvora. Tada dolazi do izražaja Peltijev efekat koji smanjuje vrednost efektivno generisanog napona čak do 36%. Simulacije su pokazale i da efekat radijacionog zračenja uvećava vrednost generisanog napona do 30%. Uočava se da keramički hladnjaci, bilo čvrsti ili mikroporozni, pokazuju najbolje performanse za data geometrijska ograničenja. Rezultati analize mogu potpuno analogno da se primene i na mini sisteme za hlađenje/grejanje.



Sl. 6. Zavisnost napona na potrošaču od temperaturne razlike ΔT=T_{hot}-T_{amb} za sklop sa HS4. Simboli i pune linije – eksperiment; zvezdice i isprekidane linije - simulacije.

ZAHVALNICA

Ovaj rad je realizovan u okviru projekta finansiranog od strane Ministarstva za prosvetu, nauku i tehnološki razvoj Republike Srbije (Ev. br. 451-03-9/2021-14/200102).

LITERATURA

[1] H. J. Goldsmid,."Thermoelectric Modules and Their Application", in *Introduction to Thermoelectricity. Springer Series in Materials Science*, vol 121. pp. 197-220, Springer, Berlin, Heidelberg, 2016.

[2] N. Jaziri, A. Boughamoura, J. Müller, B. Mezghani, F. Tounsi, M. Ismail, „A comprehensive review of Thermoelectric Generators: Technologies and common applications“, *Energy Reports*, vol. 6, Suppl. 7, pp. 264-287, 2020.

[3] M. Baldry, V. Timchenko, C. Menictas, „Optimal design of a natural convection heat sink for small thermoelectric cooling modules“, *Appl. Therm. Eng.*, vol. 160, 114062, Sept., 2019.

[4] European Thermodynamics Ltd, https://www.europeanthermodynamics.com/

[5] AMEC Thermasol, https://amecthermasol.co.uk/datasheets/MPC%20Ceramic%20Series.pdf

[6] Versarien Technologies Ltd, http://www.versarien-technologies.co.uk/.

[7] FreeCAD 3D parametric modeler, https://www.freecadweb.org.

[8] ANSYS Inc., Multiphysics simulation, https://www.ansys.com/products/platform/multiphysics-simulation.

[9] A. Sarkar, S. Mahapatra, "Role of surface radiation on the functionality of thermoelectric cooler with heat sink", *Appl. Therm. Eng.*, vol. 69, no. 1-2, pp. 39-45, Aug. 2014.

[10] M. Liao, Z. He, C. Jiang, X. Fan, Y. Li, F. Qi, „A three-dimensional model for thermoelectric generator and the influence of Peltier effect on the performance and heat transfer“, *Appl. Therm. Eng.*, vol. 133, pp. 493-500, March, 2018.

[11] J. Wang, J. Carson, M. North, D. Cleland, „A new approach to modelling the effective thermal conductivity of heterogeneous materials“, *Int. J. Heat Mass Transf.*, vol. 49, no. 17-18, pp. 3075-3083, Aug. 2006.

[12] N. Nguyen, K. Pochiraju, „Behavior of thermoelectric generators exposed to transient heat sources“, Appl. Thermal Eng., vol. 51, no. 1, pp.1-9, Mar., 2013.

[13] M. Marjanović, A. Prijić, B. Randjelović, Z. Prijić, „A transient modeling of the thermoelectric generators for application in wireless sensor network nodes“, *Electronics-MDPI*, vol. 9, 1015, Jun. 2020.

[14] A. Prijić, M. Marjanović, Lj. Vračar, D. Danković, D. Milić, Z. Prijić, „A Steady-State SPICE Modeling of the Thermoelectric Wireless Sensor Network Node“, Proc. of the 4th IcETRAN, Kladovo, Serbia, pp. MOI2.3.1-6, June 05-08, 2017.

ABSTRACT

Thermoelectric module-heatsink assemblies are widely used in systems converting thermal into electrical energy and vice versa. In this paper, performances of the selected thermoelectric module in conjunction with six different heatsinks of similar dimensions under natural cooling conditions are analyzed. The efficiency of the assemblies is determined by the value of the voltage transferred to the electric load at different temperature loads. Experimental measurements and appropriate numerical simulations were performed, and the contributions of individual thermoelectric effects and heat dissipation mechanisms were analyzed. It is shown that flat heatsinks made of microporous materials can adequately replace extruded aluminum heatsinks in energy harvesting systems under natural cooling conditions.

Performances of the thermoelectric module-heatsink assemblies for energy harvesting under natural cooling

Aleksandra Stojković, Miloš Marjanović, Jana Vračar, Aneta Prijić, Zoran Prijić

Analiza uporednog praćenja temperature površine ohlađenih materijala pri njihovom zagrevanju do ambijentalne temperature

Stevan Đenadić, Ljubiša Tomić, Vesna Damnjanović i Katarina Nestorović

***Apstrakt*—U datom eksperimentalnom radu su termografskim metodama ispitivane termofizičke karakteristike različitih materijala. Za ispitivanja su izabrani mesingani niovčić, ugalj i pelet, jer imaju različita termoizolaciona svojstva i različito stanje površine. Uzorci su najpre hlađeni, a zatim je njihovo zagrevanje na sobnoj temperaturi praćeno termalnom kamerom. U drugom delu eksperimenta, korišćena je impulsna termografija, a zagrevanje površine uzoraka vršeno je svetlosnom pobudom.**

***Ključne reči*—termografija; impulsna termografija; nedestruktivno testiranje materijala; mesing; ugalj; pelet.**

I. Uvod

Metode infracrvene termografije koje se koriste za otkrivanje zagrejanih tela, beskontaktnu procenu temperature površine objekata i nedestruktivno testiranje materijala, danas imaju široku primenu ne samo u bezbednosnom sektoru već i u mnogobrojnim privrednim i naučnim granama [1,2]. Još neke mogućnosti primene termografskih metoda odnose se na praćenje procesa neželjenog zapaljenja pojedinih gasova, tečnosti, lako zapaljivih čvrstih supstanci, ali i na praćenje procesa sagorevanja materijala različitih kalorijskih vrednosti [3-6]. Širokoj rasprostranjenosti termografskih metoda, doprinele su sve pristupačnije cene odgovarajuće komercijalne opreme.

II. Test uzorci

Kao uzorak broj 1 izabran je novčić od pet dinara [7]. Na njegovoj prednjoj strani je brojem i slovima oznaka nominalne vrednosti, a na njegovoj zadnjoj strani je reljef manastira Krušedol i oznaka godine kovanja. Prečnik novčića je 24 mm a masa 5,78 g. Novčić je višeslojni. Jezgro je izrađeno od legure niskougljeničnog čelika. Obostrano je dvostruko presvučen galvanskom prevlakom, i to prvim slojem bakra (do jezgra), i drugim slojem mesinga (na površini). Mesing kao legura bakra i cinka široko se koristi ne samo u kovnici novca za prevlaku apoena, već i u elektrotehnici, kao konstruktivni ukrasni materijal, za izradu muzičkih instrumenata, i dr. Mesing koji se koristi za kovanice (oznaka: CuZn49) sadrži 65% bakra i 35% cinka. Gustina mu je 8,7 g/cm^3, specifični toplotni kapacitet 380 J/(kg ^{o}C), toplotni kapacitet po jediničnoj zapremini 3,2 J/(cm^3 ^{o}C), toplotna provodnost 130 W/(m ^{o}C), koeficijent toplotne difuzije 32×10^{-6} m^2/s, a poseduje i antibakterijsko dejstvo. Mesingana legura koja je na površini novčića je termoprovodni materijal, neravna je ali je reljef pravilnih oblika.

Kao uzorak broj 2, izabran je pljevaljski ugalj [3], koji spada u kategoriju visokokaloričnog mrkog lignita. Njegova kalorijska vrednost (toplotna moć) se kreće u granicama od 15424 kJ/kg do 14021 kJ/kg. Ugalj je po svojoj prirodi nehomogeni materijal. Ovu vrstu uglja karakteriše izuzetno mali sadržaj sumpora, nizak procenat vlage i pepela, a granulacija ove vrste uglja je od 40 mm do 80 mm. Masa uzorka iznosila je oko 3,5 g. Ugalj je termoizolacioni materijal a njegova površina je hrapava.

Za uzorak broj 3, izabran je bukov pelet [6]. Pelet je čvrsto gorivo u vidu briketa koje se dobija u specijalnim presama koje rade na principu kompresije prethodno prerađenog materijala. Materijal se sastoji od 70% tvrdog drveta i 30% mekog drveta. U procesu proizvodnje, drvo se najpre usitnjava, zatim se separira u krupnu i sitnu piljevinu (brašno) i skladišti u velike silose, odakle se kao pripremljena drvena smeša presuje u pelet. Briketi su valjkastog oblika, visina valjka je do 4 cm, a prečnik do 6,1 mm. Gustina bukovog peleta je 6,36 g/cm^3, toplotna moć je 17300 kJ/kg a masa po briketu je približno 2,97 g. Pelet je termoizolacioni materijal a površina briketa je glatka.

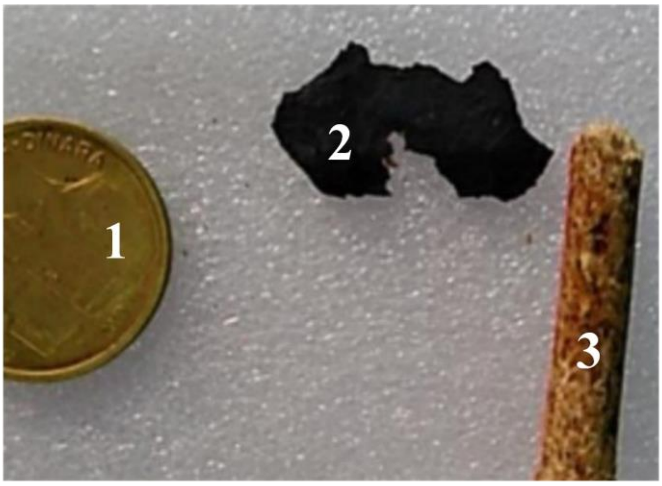
III. Eksperimentalna postavka

U prvom delu eksperimenta primenjena je metoda pasivne termografije. Sva tri uzorka, najpre su postavljena na izolacioni materijal od plastike (fotografija prikazana na slici 1.) a zatim hlađena zajedno sa njim u zamrzivaču. Nakon jednog sata, uzorci na plastičnoj podlozi izneti su iz zamrzivača i postavljeni na sto za snimanje termalnom kamerom FLIR S60.

Neposredno nakon hlađenja, u nastavku eksperimenta, snimljena je serija termograma u odgovarajućim vremenskim intervalima, sa ciljem praćenja zagrevanja ohlađenih uzoraka, različitih termofizičkih svojstava sve do ambijentalne

Stevan Đenadić–Univerzitet u Beogradu, Rudarsko-geološki fakultet, Đušina 7, 11000 Beograd, Srbija (e-mail: stevan.djenadic@rgf.bg.ac.rs)
Ljubiša Tomić–Vojnotehnički institute, Ratka Resanovića 1, 11030 Beograd, Srbija (e-mail: ljubisa.tomic@gmail.com)
Vesna Damnjanović–Univerzitet u Beogradu, Rudarsko-geološki fakultet, Đušina 7, 11000 Beograd, Srbija (e-mail: vesna.damnjanovic@rgf.bg.ac.rs)
Katarina Nestorović–Vojnotehnički institute, Ratka Resanovića 1, 11030 Beograd, Srbija (e-mail: katarina.nestorovic933@gmail.com)

temperature. Temperatura ambijenta, mereno termografski, iznosila je 21 ^{o}C.



Sl. 1. Fotografija uzoraka postavljenih na termoizolacionom materijalu od plastike

U drugom delu eksperimenta, za ispitivanja je primenjena metoda impulsna termografija (PT – *Pulse thermography*), koja spada u metode aktivne termografije. Pobuda uzoraka je impulsna - svetlosnim fluksom u kratkom vremenskom intervalu. Naime, u eksperimentalnoj postavci sada je, pored termalne kamere FLIR S60, kao impulsni svetlosni izvor za osvetljavanje uzoraka (za zagrevanje njihovih površina) korišćen fotografski blic YASHICA CS-250AF. Trajanje i intenzitet svetlosnog impulsa je kalibrisano [8]. Dakle, uzorci nisu bili hlađeni, ali je jedan od uzoraka (briket od peleta), pre snimanja, jednim delom, bio neko vreme potopljen u vodi. Briket peleta je i posle kraćeg sušenja bio po dužini homogeno vlažan.

IV. Rezultati i diskusija

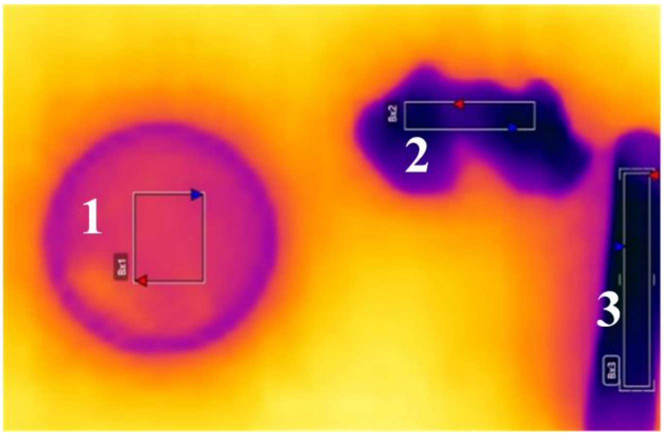
Eksperiment je osmišljen tako da se "uporedo" pod "istim uslovima" prate promene temperatura površina ohlađenih materijala sve do njihovog zagrevanja do ambijentalne temperature. Cilj je bio da se utvrdi brzina promene temperature (zagrevanja) u zavisnosti od termofizičkih svojstava uzoraka, sa ciljem da se potom izvrši procesiranje snimljenih termograma u infracrvenom programskom paketu ThermaCAM Researcher 2.9 i analiza praćenja zagrevanja.

Na slici 2, prikazan je termogram iste scene kao i na slici 1 (u vidljivoj oblasti elektromagnetskog spektra), nastale u istom trenutku kao i termogram u infracrvenoj oblasti elektromagnetnog spektra (8–14 μm). Termogram prikazan na slici 2 snimljen je neposredno posle iznošenja iz zamrzivača (oko 1 minut kasnije), nakon čega su, u nastavku eksperimenta, snimljeni termogrami u određenim vremenskim intervalima sa ciljem praćenja "zagrevanja" ohlađenih uzoraka sve do ambijentalne temperature. Uzorci su bili izloženi početnim stacionarnim uslovima dovoljno dugo da dođe do uspostavljanja termičke ravnoteže. Naglim pomeranjem uzoraka iz stacionarnog temperaturnog stanja (pothlađenog) u odnosu na ambijentalnu temperaturu, tj. sa niže na višu ambijentalnu temperaturu, dolazi do zagrevanja uzoraka.

Procesi zagrevanje uzoraka nisu identični, iako su isti eksperimentalni uslovi, usled njihovih različitih termofizičkih svojstava. Na brzinu zagrevanja uzoraka uticali su različiti faktori - pre svega njihova fizička svojstva, ali i parazitni toplotni izvori, čiji je uticaj više dolazio do izražaja kod glatkih (refleksivnih) površina. Naime novčić ima reljefnu ali refleksivnu površinu, pa parazitna refleksija njegove površine stvara lažnu sliku o pravoj radijacionoj temperaturi njegove površine. Dok kod uzorka sa višim koeficijentom emisivnosti a niskim koeficijentom refleksije (tj. kod ugalja i briketa od peleta) parazitna refleksija slabije utiče na radijacionu temperaturu površine.

Na termogramu predstavljenom na slici 2, radi termografske procene njihove radijacione temperature, markirane su tri oblasti označene kao: Bx1, Bx2, i Bx3. Prosečna temperatura uzorka 1 (novčić) iznosila je 7,2 ^{o}C. Masa novčića znatno je različita od masa uzorka peleta i uzorka uglja. Prosečna temperatura uzorka 2 (ugalj) iznosila je - 3,6 ^{o}C. Prosečna temperatura uzorka 3 (pelet) iznosila je - 4,4 ^{o}C.

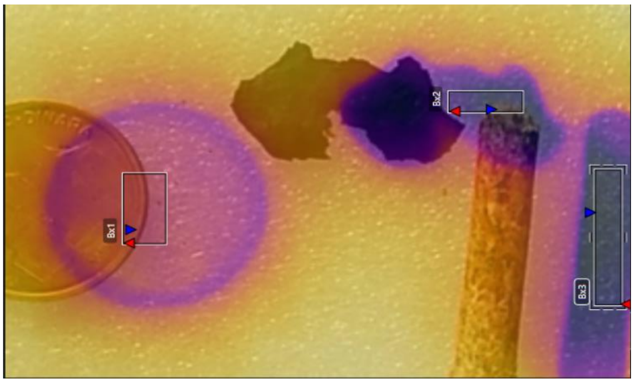
Na istom termogramu, uočljiva je i refleksija toplih predmeta u okolini koja je dominantna na površini visokorefleksionog materijala (novčića) prekrivenog slojem blage kondenzacije vlage iz okolnog vazduha na hladnoj površini.



Sl. 2. Termogram uzoraka postavljenih na izolacioni materijal od plastike (ista scena kao na Sl. 1.)

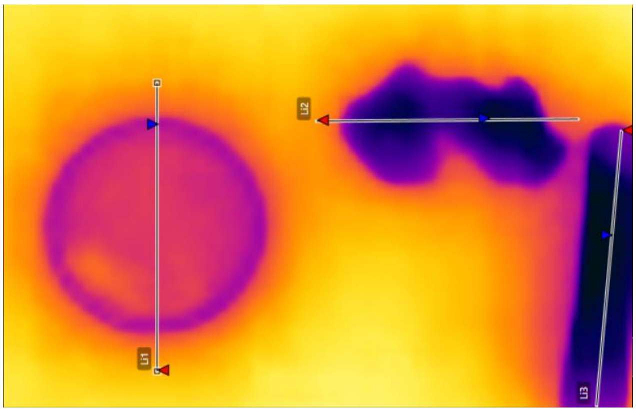
Ispitivanja na osnovu snimljenih termograma ukazuju da se ponašanje ohlađenih uzoraka različitih termofizičkih svojstava, u odnosu na topliju okolinu ambijenta, tokom vremena odvija svojstveno fizičkoj prirodi materijala.

Na slici 3 prikazana je fotografija uzoraka u formatu *thermal blending* zajedno sa termogramom uzoraka u istoj sceni. Istovremeno su prikazane i slika snimljena video kamerom i termogram snimljen termalnom kamerom, jer se radi o različitim objektivima koji su pomereni prostorno za oko 1 cm od centra do centra otvora jednog i drugog objektiva. Upravo iz tog razloga su u horizontalnoj ravni pomerene slike uzoraka (rastojanja jednog i drugog sočiva objektiva od ravni scene).



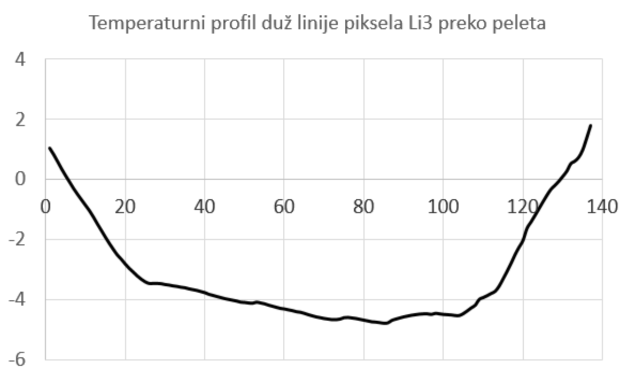
Sl. 3. Fotografija (*format thermal blending*) uzoraka postavljenih na plastičnom izolacionom materijalu i termogram iste scene

Na slici 4 prikazan je termogram (flir_20210527T15:18:30) uzoraka postavljenih na plastičnom izolacionom materijalu na kome su povučene marker linije radi određivanja temperaturnog profila, obeležene kao Li1-novčić, Li2-ugalj, i Li3-pelet.



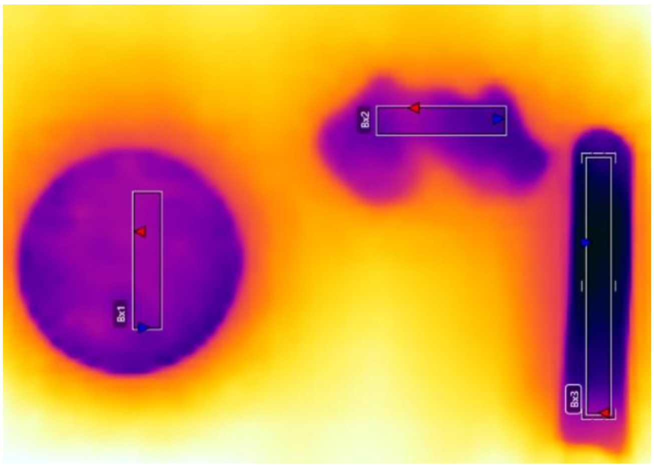
Sl. 4. Termogram uzoraka postavljenih na plastičnom materijalu na kome su povučene marker linije radi određivanja temperaturnog profila

Temperaturni profil duž marker linija Li3 povučene preko briketa peleta prikazan je na slici broj 5.



Sl. 5. Temperaturni profil duž marker linije Li3 (na Sl. 4) ohlađenog uzorka briketa peleta snimljen na početku zagrevanja na sobnoj temperaturi

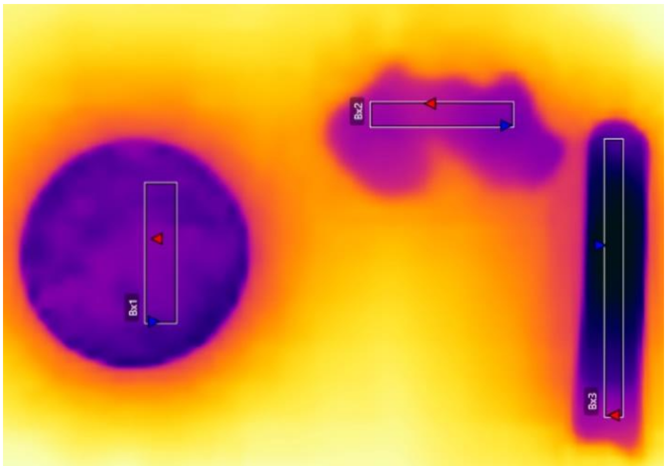
Na slici 6 prikazan je termogram (flir_20210527T15:19:23) snimljen 53 s nakon iznošenja uzoraka iz zamrzivača. Sa navedenog termograma se očitava da je srednja temperature mesinganog novčića 10,7 °C, a temperature komada uglja pljevlja 10,0 °C, što je za 0,7 °C niže od temperature novčića. Dakle, uočava se da je temperatura uglja niža od temperature mesinganog novčića.



Sl. 6. Termogram uzoraka snimljen 53 s nakon hlađenja

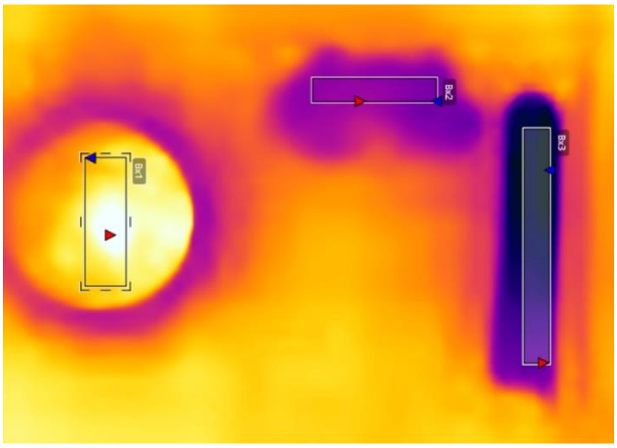
Na slici 7 prikazan je termogram (flir_20210527T15:20:06) snimljen 96 s kasnije nakon hlađenja, tj. 43 s kasnije u odnosu na termogram prikazan na slici 6. Sa navedenog termograma se očitava da srednja temperature mesinganog novčića iznosi 13,2 °C, a temperature komada uglja 14,4 °C, što je za 1,2 °C više od temperature novčića.

Uočava se da je sada prikazan „obrnut slučaj“, tj. temperatura uglja je viša od temperature mesinganog novčića. Dakle, može da se zaključi da zbog visokorefleksivne mesingane prevlake novčić sporije prima toplotu od okoline. Viši koeficijent emisivnosti ima ugalj pljevlja u odnosu na novčić presvučen mesinganom prevlakom.



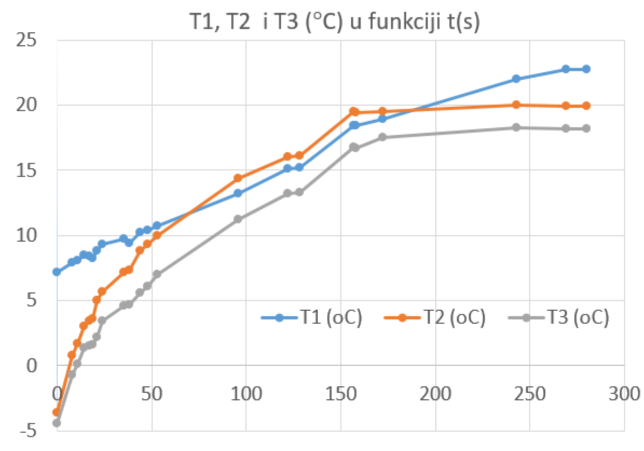
Sl. 7. Termogram uzoraka snimljen 96 s nakon hlađenja

Očigledno da od oko 70-te sekunde od početka zagrevanja, pa do 170-te sekunde, dolazi do isparavanja filma vodene pare na novčiću i sporije apsorpcije toplote iz okoline u odnosu na ugalj. Na slici 8, termogram (flir_20210527T15:23:44) nastao 280 s nakon hlađenja, vodena para je isparila sa njegove površine (novčić se osušio), pa do izražaja dolazi (dominantna je) njegova visokorefleksivna prevlaka na površini, u odnosu na parazitnu refleksiju zračenja toplih predmeta iz okoline scene (na primer tople ruke ekperimentatora).



Sl. 8. Termogram nastao 280 s nakon hlađenja

Na slici 9 predstavljen je grafik zavisnosti temperature od vremena pri zagrevanju hlađenog uzoraka: novčića T_1, uglja T_2 i peleta T_3. Početne temperature uzorka uglja i peleta su – 3,6°C i – 4,4°C respektivno – radi se o toplotnim izolatorima sa visokim vrednostima emisivnosti. Početna temperaturu novčića je 7,2 °C – radi se o toplotnom provodniku visokorefleksivne površine pa na njega veliki uticaj ima parazitno zračenja okoline.



Sl. 9. Zavisnost temperature od vremena pri zagrevanju hlađenog uzorka: novčić – plava serija T1; ugalj – narandžasta serija T2; i peleta – siva serija T3

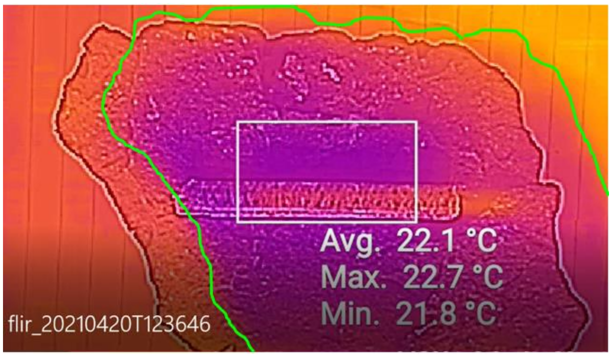
U tabeli broj 1 navedeni su rezultati termografskog merenja temperature tokom vremena u procesu zagrevanja hlađenih uzoraka 1, 2, i 3 do sobne temperature.

TABLE I

RAST TEMPERATURE TOKOM VREMENA U PROCESU ZAGREVANJA HLAĐENIH UZORAKA (NOVČIĆ – T_1; UGALJ – T_2; I PELET – T_3)

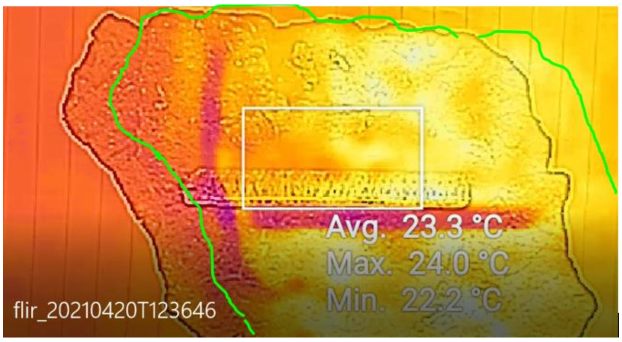
	t (s)	T_1 (ºC)	T_2 (ºC)	T_3 (ºC)
flir_20210527T15:18:30	0	7.2	-3.6	-4.4
flir_20210527T15:18:38	8	7.9	0.8	-0.7
flir_20210527T15:18:41	11	8.1	1.7	0.1
flir_20210527T15:18:44	14	8.5	3	1.4
flir_20210527T15:18:47	17	8.4	3.4	1.5
flir_20210527T15:18:49	19	8.2	3.6	1.6
flir_20210527T15:18:51	21	8.8	5	2.2
flir_20210527T15:18:54	24	9.3	5.7	3.4
flir_20210527T15:19:05	35	9.7	7.2	4.6
flir_20210527T15:19:08	38	9.4	7.3	4.7
flir_20210527T15:19:14	44	10.2	8.8	5.6
flir_20210527T15:19:18	48	10.4	9.3	6.1
flir_20210527T15:19:23	53	10.7	10	7
flir_20210527T15:20:06	96	13.2	14.4	11.2
flir_20210527T15:20:32	122	15.1	16	13.2
flir_20210527T15:20:38	128	15.2	16.1	13.3
flir_20210527T15:22:07	157	18.4	19.5	16.8
flir_20210527T15:22:08	158	18.4	19.4	16.7
flir_20210527T15:22:22	172	18.9	19.5	17.5
flir_20210527T15:23:07	243	22	20	18.3
flir_20210527T15:23:33	269	22.7	19.9	18.2
flir_20210527T15:23:44	280	22.7	19.9	18.2

U drugom delu eksperimenta, u merenjima je primenjena PT termografija. Na slici 10 prikazan je termogram površine drvenog peleta i uglja dok su u termodinamičkoj ravnoteži, tj. neposredno “pre” okidanja blica.



Sl. 10. Termogram površine drvenog peleta i uglja neposredno pre okidanja blica

Na slici 11 prikazan je termogram površine peleta i uglja pola sekunde nakon osvetljavanja blicem, tj. nakon primanja toplotnog impulsa. Nakon okidanja blica dolazi do generisanja površinski homogenog toplotnog impulsa, koji se procesom toplotne difuzije prostire kroz materijale, usled čega dolazi do temperaturne perturbacije na površini uzoraka. Toplotni impuls koji se formirao na nehomogenoj površini uglja, prostire se kroz uzorak različito, u zavisnosti od sastava uglja, to jest koeficijenta difuzivnosti, na primer čistog uglja i uglja sa više primesa gline. Toplije i hladnije oblasti na površini uglja su nastale usled heterogenog sastava uglja (više i manje gline), razlike u emisivnosti oblasti sa glinom i čistog uglja, gde je apsorbovano više energije od svetlosnog impulsa. Pelet je homogenije zagrejan pošto je kompaktnijeg od uglja koji je „ljuspastog“ sastava. Deo potapanog briketa peleta, iako ponovo suv, bio je hrapaviji od dela peleta koji nije potapan u vodu (tj. hrapavije površine), što znači da mu je hrapavost površine degradirana trajno pa ima promenjenu emisivnost. Na termogramu se za uzorak 3 (pelet) u donjem delu uočava svetlija površina koja odgovara delu uzorka koji je bilo potopljena u vodu kratko vreme, nakon čega se osušio ali sa degradacijom koeficijenta emisivnosti u odnosu na nepokvašeni gornji deo koji nije bio u dodiru sa vodom.



Sl. 11. Termogram površine drvenog peleta i uglja pola sekunde nakon osvetljavanja blicem

Dati rad je pionir ideje da se utvrdi može li se termografija koristi i za grubu procenu kalorijske vrednosti uglja, zasnovanoj na vremenskoj promeni površinske temperature. Inače, utvrđivanje kalorijske vrednosti uglja, koji u termoelektranama Srbije pokriva oko 80% proizvodnje električne energije, je izuzetno složen proces. Propisan je ISO standardom B.H8 318:1972 koji dozvoljava mernu nesigurnost u proceni čak iznad 20 % [9]. Vrednost toplotnog kapaciteta uzorka jeste povezana sa promenom temperature površine, ali da bi se utvrdila veza i sa njegovom kalorijskom vrednošću, treba izvršiti merenja sa uzorcima potpuno istih masa i oblika, uz uzimanje u obzir svih navedenih uticaja u datom eksperimentu, ali i njihovu minimizaciju kroz dizajn eksperimenta.

V. Zaključak

Rezultati predstavljeni u radu, dobijeni na osnovu analize termograma pri proučavanju različitih materijala kako pasivnom tako i aktivnom impulsnom termografijom sa ciljem procene njihovih termofizičkih svojstava. Pri istim uslovima, snimani su uzorci sačinjeni od materijala potpuno različitih termofizičkih karakteristika, i to mesingani novčić, ugalj i pelet. Navedeni materijali su izabrani jer imaju različita termoizolaciona svojstava i različito stanje površine. Termogrami nastali pasivnom metodom, tokom procesa prirodnog zagrevanja (na sobnoj temperaturi), prethodno ohlađenih uzoraka od različitih materijala (termoprovodnih i termoizolacionih) ukazuje na to da termografske metode mogu biti alternativna metoda u proceni termofizičkih svojstava materijala. Problemi koji javljaju pri praćenju promena temperature površina uzoraka kada se primenjuje metoda aktivne impulsne termografije nastaju usled nehomogene strukture (konkretno, naglašene nehomogenosti sastava uglja, hrapavosti površine i nepravilnog oblika) mogu biti svedeni na minimum izborom homogenih, pravilno obrađenih komada (uzoraka) istih oblika.

Zahvalnica

Rad je nastao uz podršku Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije - kontakt broj: 451-03-9/2021-14/200126.

Literatura

[1] X. P. Maldague, *Theory and Practice of Infrared Technology for Nondestructive Testing*, John Wiley & Sons, New York, USA, 2001.
[2] G. Gaussorgues, *Infrared Thermography*. English language edition. Springer-Science, Berlin, Germany, 1994.
[3] S. Đenadić, V. Damnjanović, P. Jovančić, D. Ignjatović, D. Jovković, F. Miletić, “Primena termografije u rudarstvu”, IX Međunarodna konferencija Ugalj 2019, Zlatibor, Srbija, s. 57-66, 23-26 Oktobar 2019.
[4] S. Đendić, Lj. Tomić, V. Damnjanović, P. Jovančić, D. Jovković, „Testing the energy value of different types of coal by the method of Active Thermography”, 9th International scientific conference on defensive technologies OTEH 2020, Belgrade, Serbia, pp. 424-428, 15-16 October 2020.
[5] Lj. Tomić; M. Milinović: Experimental research of limits for thermal modulation transfer function,*Thermal Science*, vol. 13, no. 4, pp. 119 – 128, 2009.
[6] Ljubiša Tomić., Vesna Damnjanović., Katarina Mišković., Boban Bondžulić., Dragan Knežević., Aleksandar Kovačević. (2018) Analiza uticaja temperature materijala na primenu metode impulsne termografije. *ETRAN 2018* (*62. konferencija za elektroniku, telekomunikacije, računarstvo, automatiku i nuklearnu fiziku*). *Zbornik radova*, MO. s. 314-318, 11-14. jun, Palić, Srbija.
[7] Народна банка Србије www.nbs.rs/sr_RS/novac-i-p
[8] Lj. D. Tomić, „Nondestructive evaluation of the thermophysics properties materials by IR thermography“, Ph.D. dissertation, School of Electrical Engineering, University of Belgrade, Serbia, 2013.
[9] *Solid mineral fuels – Determination of gross calorific value by bomb calorimetric method and calculation of net calorific value*, SRPS ISO 1928:2015 https://iss.rs/en/project/show/iss:proj:47873

Abstract

In the given experimental work, thermophysical characteristics of different materials were examined by thermographic methods. Brass coin, coal, and pellets were chosen for the tests - because they have different thermal insulation properties and different surface condition. The samples were first cooled and then let to be warmed to room - temperature was continuously being recorded by a thermal camera. In the second part of the experiment, pulsed thermography was used, and the surface of the samples was heated by light excitation.

Analysis of comparative monitoring of surface temperature of cooled materials during their heating to ambient temperature

Stevan Djenadic, Ljubisa Tomic, Vesna Damnjanovic, Katarina Nestorovic

Synthesis and characterization of thin copper coatings obtained by sonoelectrodeposition method

Ivana Mladenović, Jelena Lamovec, Stevan Andrić, Miloš Vorkapić, Marko Obradov, DanaVasiljević-Radović, Vesna Radojević and Nebojša D. Nikolić

***Abstract*—Influence of an intensity of ultrasonic mixing of electrolyte in a temperature range of 27–37 °C and ultrasonic power intensity in the range of 3.77–18.84 W/cm^2 (10–50 %) on a synthesis of fine-grained copper deposits was examined. Copper coatings were electrodeposited on a brass substrate in direct current (DC) regime with an applied current density of 50 mA·cm^{-2}. The laboratory-made copper sulfate electrolyte was used without or with addition of additives. The variation of temperature under sonoelectrodeposition process and variation mixing intensity of electrolyte were ensured by using an ultrasonic probe. The produced Cu coatings were examined by optical microscope (OM) in order to observe the microstructural modification with variation ultrasonic parameters and for measuring imprints of Vickers indenter. The micro hardness properties of composite systems were characterized using Vickers micro indentation test. The composite hardness models Chicot-Lesage and Chen-Gao were used for the determination the coatings hardness and adhesion evaluation. Application of Atomic Force Microscopy (AFM) technique also confirmed the strong influence of ultrasonic mixing conditions of electrolyte onto change of the microstructure of copper deposits and surface roughness of the coatings. The maximum hardness, good adhesion properties and minimum micro surface roughness was obtained for the fine-grained Cu coating produced with amplitude of 50 % ultrasonic mixing of electrolyte without additives and 30 % for electrolyte with additives.**

***Index Terms*— ultrasonic probe; microindentation; composite system; coatings; adhesion; sonoelectrodeposition.**

I. INTRODUCTION

THE effects of ultrasonic electrolyte mixing can be seen from a few aspects: chemical, mechanical and theirs combination. The chemical effects of ultrasound are due to the “implosion of microbubbles”, generating free hydroxyl radicals [1, 2] with high chemical reactivity, while mechanical effects are caused by “shock waves” formed during symmetric cavitation or by “microyets” formed during asymmetric cavitation [1]. The use of ultrasound in a reaction system provides specific activation based on a physical phenomenon, known as “acoustic cavitation”. Cavitation is caused by a longitudinal sound wave, which results in a change in local pressures and temperatures in the liquid electrolyte [2]. The change in pressure causes the gas bubbles to form and burst, which contributes to electrolyte mixing [3]. The scheme of the mechanism of formation, growth and bursting of bubbles under ultrasound conditions are shown in Fig. 1.

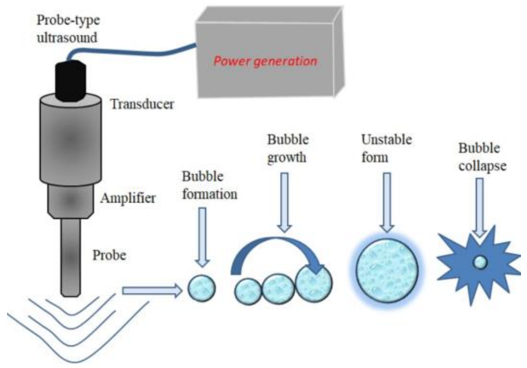


Fig. 1. Schematic presentation of bubble implosion under ultrasonic conditions in electrolyte near cathode surface [3].

Electrodeposition in presence of ultrasound (US-ED) may have effects in terms of: degradation of the reaction, enhanced electrochemical diffusion processes and mass-transport, increase electrochemical rates and current efficiencies, decrease electrode overpotentials [4]. The effects of ultrasonic mixing of electrolyte and influence on the metallic coatings properties have also reported in the literature: change morphology (reduction in grain size or change grain orientation), influence on residual stress, that hence good adhesion and hardness, the increase of brightness has been also observed, etc [4–7]. Ultrasonic deposition (sonoelectrodeposition) is often used in combination with the co-deposition of particles as reinforcement into a metal matrix composite [8].

Using AFM technique, the coating surface roughness was investigated in function of grow rate (applied current density) [9], film thickness (deposition time), variation substrates and mixing conditions [10], as well as variation electrolytes [10, 11].

Ivana Mladenović, Stevan Andrić, Miloš Vorkapić, Marko Obradov, Dana Vasiljević-Radović and Nebojša Nikolić are with the Institute of Chemistry, Technology and Metallurgy, University of Belgrade, Njegoševa 12, 11 000 Belgrade, Serbia (e-mail: ivana@nanosys.ihtm.bg.ac.rs), (e-mail: stevan@nanosys.ihtm.bg.ac.rs), (e-mail: worcky@nanosys.ihtm.bg.ac.rs), (e-mail: marko.obradov@nanosys.ihtm.bg.ac.rs), (e-mail: dana@nanosys.ihtm.bg.ac.rs), (e-mail: nnikolic@ihtm.bg.ac.rs).

Jelena Lamovec is with the University of Criminal Investigation and Police Studies, Cara Dušana Street 196, 11 000 Belgrade, Serbia (e-mail: jelena.lamovec@kpu.edu.rs).

Vesna Radojević is with the Faculty of Technology and Metallurgy, University of Belgrade, Karnegijeva 4, 11 000 Belgrade, Serbia (e-mail: vesnar@tmf.bg.ac.rs).

Hardness and adhesion testing are the most important and widely used methods for assessing the structural and mechanical properties of the composite systems and coatings [12]. If the thickness of the coating is very small, the influence of the substrate must be considered during the coating hardness determination [13] and mathematical composite hardness models are used for this purpose [14–19].

Hence, the present experiment aims here to find the synergistic effects of ultrasound agitation, temperature oscillation, while increasing the power of ultrasonic mixing influence on the roughness, adhesion and micro hardness of copper coatings.

II. COMPOSITE HARDNESS MODELS

Multilayer complex structures such as thin coatings (films) on bulk substrates are often used in fabrication of MEMS devices. A thin coating deposited electrochemically on a conductive substrate can be considered as a composite system. Determination of the hardness of the coating, independent of the influence of the substrate hardness, is not possible in the entire region of the applied indentation loads. For this reason, composite hardness mathematical models are used to estimate the absolute hardness of a thin coating [14–19], and two adequate ones have been selected in this paper. The composite hardness model of Chicot-Lesage (C–L) was found to be suitable for experimental data analysis (for composite hardness values, H_c) and coating hardness determination (H_{coat}) based on our previous research [14–16]. Using this model, it is possible to calculate the absolute value of the coating hardness for each individual indentation load. The model is based on Meyer's law which expresses the variation of the size of the indent (d) in function of the applied load (P) [13]:

$$P = a^{*} \cdot d^{n^{*}} \tag{1}$$

The variation part of the hardness number with load is represented by the factor n^*. Chicot-Lesage adopted the following expression [16]:

$$F\left(\frac{\delta}{d}\right) = \left(\frac{\delta}{d}\right)^{m} = f \quad where\ m = \frac{1}{n^{*}} \tag{2}$$

In equation (2) m is the composite Mayer’s index, d is average diagonal size of Vickers indent and δ is coating thickness. The value of m is calculated by a linear regression performed on all the experimental data obtained for a given coating/substrate couple and deduced from the relation [13]:

$$\ln d = m \cdot \ln P + b \tag{3}$$

The composite hardness can be expressed by the following relation [13, 16]:

$$H_C = (1-f)/\left(1/H_s + f \cdot \left(\frac{1}{H_{coat}} - \frac{1}{H_s}\right)\right) + \\ + f \cdot \left(H_s + f \cdot (H_{coat} - H_s)\right) \tag{4}$$

Hardness of the film is the positive root of the next equation [13, 16]:

$$\begin{aligned} &A \cdot H_{coat}^2 + B \cdot H_{coat} + C = 0, \quad with \\ &A = f^2 \cdot (f-1) \\ &B = \left(-2 \cdot f^3 + 2 \cdot f^2 - 1\right) \cdot H_s + (1-f) \cdot H_c \\ &C = f \cdot H_c \cdot H_s + f^2 \cdot (f-1) \cdot H_s^2 \end{aligned} \tag{5}$$

With the known value of m, only the hardness of the films remains to be calculated. The values of composite hardness and substrate hardness (H_s) are the values obtained by measuring on a Vickers tester.

For the evaluation of the adhesion properties of thin films and absolute film hardness, Chen–Gao (C–G) model was chosen [17–19]. This method introduces the composite hardness as the function of the critical reduced depth beyond which the material will have no effect on the measured hardness. The critical reduced depth, represents ratio between the radius of the plastic zone beneath the indentation and indentation depth [12, 17–19]. The large value of the critical reduced depth corresponds to good adhesion, while low values indicate poor adhesion of the films. The correlation between composite hardness value H_c and the critical reduced depth b are found as [12, 17–19]:

$$H_c = H_s + \left[\frac{(n+1) \cdot \delta}{n \cdot b \cdot h}\right] \cdot (H_{coat} - H_s) \tag{6}$$

H_s and H_{coat} are the hardness of the substrate and of the film, respectively, δ is film thickness, h is indentation depth, n is the power index and b is the critical reduced depth. The convenient value of n is found to be 1.8 for “soft film on hard substrate” and n = 1.2 for “hard film on soft substrate” [12, 17–19]. It is first necessary to perform the fitting of the composite hardness depending on the depth of indentation according to the model shown in the equation (7) [18]:

$$H_c = A + B \cdot \frac{1}{h} + C \cdot \frac{1}{h^{n+1}} \tag{7}$$

A, B and C are fitting parameters, and h is indentation depth. Indentation depth can be calculated as 1/7 of diagonal size. Then coating hardness can be expressed as [18]:

$$H_{coat} = A \pm \sqrt[n]{\frac{\left[n \cdot |B| / (n+1)\right]^{n+1}}{n \cdot |C|}} \tag{8}$$

The equation used to estimate adhesion properties coating/substrate composite system has the following form [12]:

$$H_c = \frac{7 \cdot (n+1) \cdot (H_s - H_{coat})}{n \cdot b} \cdot \frac{\delta}{d} \tag{9}$$

III. Experimental

A. Preparation of electrolyte and samples

Copper was electrodeposited from an aqueous solution of copper sulfate and sulfuric acid without (electrolyte I) and with added additives (electrolyte II). The composition of the electrolytes used in experiments is given in Table 1. Three different additives were used for electrolyte II: polyethylene glycol (PEG, molecular weight 6000), Na-3-mercapto-1-propane sulfonic acid (MPSA) and NaCl, based on the suggested recipe found in the reference [20]. The volume of the electrolyte for all experiment was 100 ml. Brass foil, ½ hard, (ASTM B36, K&S Engineering), 250 µm thick was us as cathode. Pure copper foil, cylindrical shape, was used as anode (see Fig. 2).

TABLE I
Composition of the electrolytes

Electrolyte I	g/l	Electrolyte II	g/l
$CuSO_4 \cdot 5H_2O$ H_2SO_4	240 60	$CuSO_4 \cdot 5H_2O$ H_2SO_4 PEG 6000 MPSA NaCl	240 60 1 0.0015 0.1240

An ultrasonic cleaner (model HD 2200 Bandelin / Germany) of 20 kHz ± 500 Hz frequency (f_s) equipped with a standard sonotrode (probe of 13 mm tip diameter), with power (I) of 200 W in continuous mode was used for sonication and mixing of electrolyte. In general, power intensity I_i (power irradiance) represents power, I, distributed over surface area, S (sonotrode tipe surface) as shown as $I_i = I / S$ [W/cm^2]. The temperature change, T_i, of the electrolytes was observed during the operation of the ultrasound probe with the variation of the amplitude range, P_i. The temperature change of the electrolyte was measured every 5 minutes (see Table II). Electrochemical deposition was performed under DC galvanostatic mode. The current density values were maintained at 50 mA·cm^{-2}; deposition time was fixed at 15 minutes and coating thickness was calculated according to Faraday's law.

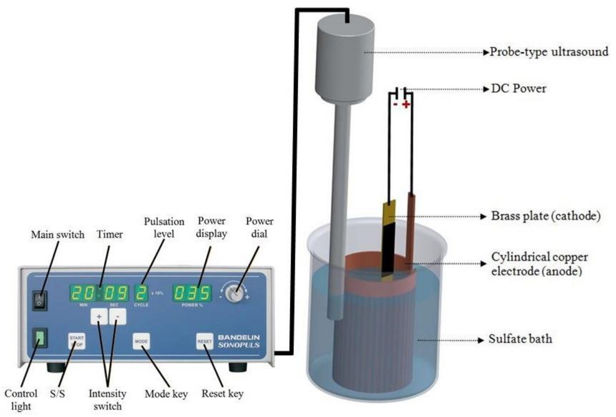


Fig. 2. Experimental setup for a sonoelectrochemical deposition process.

TABLE II
The parameters of the sonoelectrodeposition

No.	P_i / %	electrolyte	t_i / min	T_i / °C
1	10	I	1 5 10 15	28.8 29.7 30.4 32.2
3	30	I	1 5 10 15	28.3 29.6 30.6 35
5	50	I	1 5 10 15	28.9 30.9 33.2 37.2
3'	30	II	1 5 10 15	27 28.1 30.5 31.5

B. Optical microscopy

The structural properties of Cu coatings were examined by optical microscope (OM)–model Motic AE-2000 MET. A metallographic microscope–Carl Zeiss Epival Interphako was used to measure the diagonal size of Vickers indents.

C. Topography of the Cu coatings

The surface topography and roughness of the Cu coatings were examined using atomic force microscope (AFM, TM Microscopes-Vecco) in the contact mode. The values of the arithmetic average of the absolute roughness parameters (R_a) of the surface height deviation, were measured from the mean image data plane, using free software Gwyddion [21]. The values of R_a roughness parameter, calculated as average from three independent measurements at different locations of one sample of copper surface obtained by the DC regime with variation of ultrasonic intensity mixing and electrolyte composition. The scanned area was 20 µm^2 in contact mode.

D. Mechanical characterization

The mechanical properties of the composite systems were characterized using Vickers microhardness tester "Leitz, Kleinharteprufer DURIMET I" with loads ranging from 2.452 N down to 0.049 N. Three indentations were made at each indentation load from which the average diagonal and composite hardness could be calculated.

IV. Result and discussion

A. Structural characterization of the Cu coatings

Fig. 3 shows morphologies of the Cu coatings obtained using sonoelectrodeposition method with ultrasonic amplitude at 10 % (Fig. 3a), 30 % (Fig. 3b), 50 % (Fig. 3c) from electrolyte I.

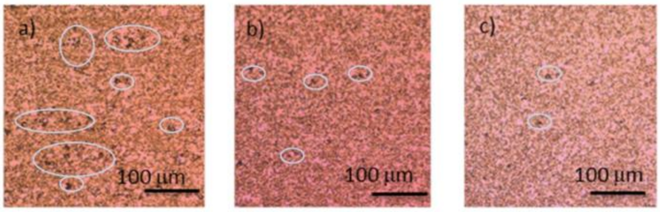


Fig. 3. Microscopic images of Cu coatings obtained in sonoelectrodeposition regime on a brass substrate from electrolyte I. The sonoelectrodeposition parameters are: j = 50 mA·cm^{-2}, t = 15 min, δ = 16.61 µm, T_{av} = 31.6 °C and f_s = 20 kHz with variation of the amplitude range of ultrasonic mixing: a) P_i = 10 %, b) P_i = 30 % and c) P_i = 50 %.

The smaller number of different size of cavitation holes were observed (Fig. 3a). It has been observed that the number of holes decreases with increasing intensity of ultrasonic mixing of the electrolyte (Figs. 3b and 3c). The bursting of large gas bubbles caused by the mixing of electrolytes at the low intensity, leads to the formation of micro cavities on the surface of the coating (Fig. 1).

B. The roughness analysis of the Cu coatings

The AFM surface areas of Cu coatings produced under different conditions are shown in Fig. 4. The average values of the roughness parameters (R_a) obtained by application of *Gwyddion* free software are given in Table III. The average roughness can be calculated as:

$$R_a = \frac{1}{N_x \cdot N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \left| z(i,j) - z_{mean} \right| \tag{10}$$

where N_x and N_y are the number of scaning points on the x-axis and y-axis; $z(i, j)$ is the height of the (i, j) measuring point, z_{mean} is the mean hight of all measuring points [22].

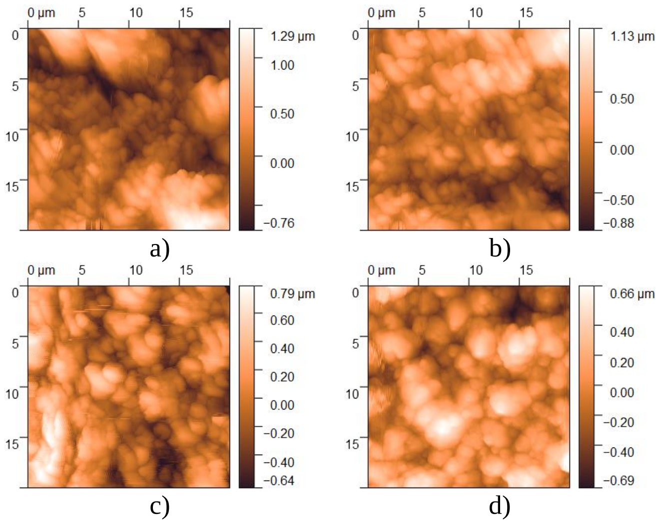


Fig. 4. 2D-AFM images of copper deposits prepared on brass substrate obtained in DC-US regime. The sonoelectrodeposition parameters are: j = 50 mA·cm^{-2}, t = 15 min, δ = 16.61 µm, T_{av} = 31.6 °C and f_s = 20 kHz with variation of the amplitude range of ultrasonic mixing: a) P_i = 10 %, b) P_i = 30 % and c) P_i = 50 % for electrolyte I and d) P_i = 30 % for electrolyte II.

Based on the results in Table III and Fig.4, a decrease in the roughness of the Cu coating from electrolyte I was observed with increasing mixing intensity (Fig. 4a–c).

The copper coating obtained from the electrolyte with additives and 30 % applied ultrasound amplitude hads minimal roughness (see Fig. 4d).

TABLE III

THE VALUES OF ROUGHNESS PARAMETER FOR SONOELECTRODEPOSITION COPPER COATINGS WITH VARIATION ULTRASONIC AMPLITUDE AND ELECTROLYTE

No.	P_i %	electrolyte	R_a / nm
1	10 %	I	270
3	30 %	I	230
5	50 %	I	180
3’	30 %	II	148.8

C. Composite hardness of copper coatings/brass systems.

The average values of the indent diagonal d (in µm), were calculated from several independent measurements on every specimen for different applied loads P (in N). The absolute substrate hardness and composite hardness values, H_c (in GPa) were calculated using the formula [23]:

$$H_c = 0.01854 \cdot P \cdot d^{-2} \tag{11}$$

where 0.01854 is a constant, geometrical factor for the Vickers indenter.

Fig. 5 shows the variation in composite hardness values with relative indentation depth (h/δ =RID), where h is indentation depth (h = d/7) for copper coatings deposited from electrolyte I with change applied ultrasonic amplitude percent.

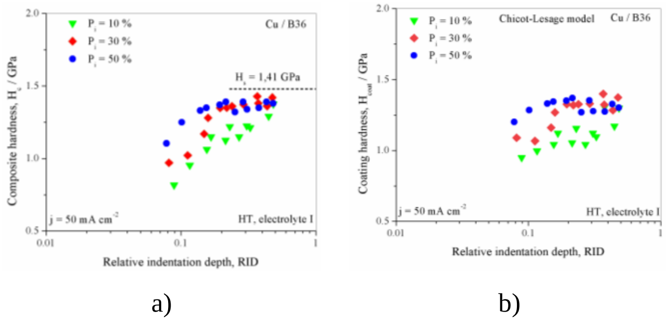


Fig. 5. Variation hardness vs. relative indentation depth for the copper coatings deposited on brass substrates with variation ultrasonic mixing of electrolyte: a) composite hardness and b) hardness of the coating calculated according to Chicot-Lesage model. The current density and deposition time were 50 mA·cm^{-2} and 15 min. Electrolyte I was used.

The effect of changing the intensity of the ultrasonic mixing of the electrolytes is reflected in the change values of the composite system hardness and the coating hardness. An increase in composite hardness with increasing electrolyte mixing was observed. The maximum value of the composite hardness was obtained for the copper coating deposited with the 50 % ultrasound mixing electrolyte (Fig. 5a) and the minimum value for the 10 %. The results of the calculated coating microhardness according to the C-L model more clearly indicate the influence of the applied ultrasound mixing condition on the microstructure and the hardness of the coating, and the highest coating microhardness value was obtained for the 50 % ultrasound amplitude settings (Fig. 5b). For shallow indentation penetration depth ($0.1 \leq$ RID ≤ 1), it was found that the response was that of the coating and substrate together. Based on the results from: Fig. 3c), Table III (sample 5), Fig. 4c and Fig. 5a,b conclude that a 50 % ultrasound amplitude is an optimal condition for copper ultrasonic deposition from electrolyte I.

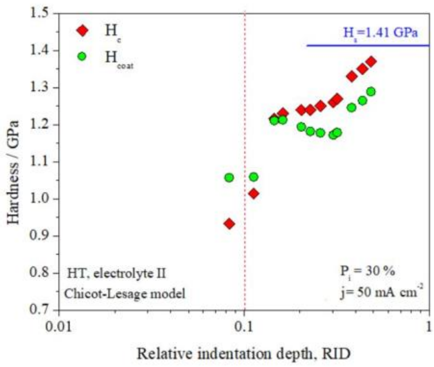


Fig. 6. Variation hardness vs. relative indentation depth for the copper coatings deposited on brass substrates with applied 30 % amplitude of ultrasonic mixing of electrolyte II. Red points are composite hardness values and green points are values for hardness of the coating calculated according Chicot-Lesage model. The current density and deposition time were: 50 mA·cm^{-2} and 15 min.

The composite hardness response has a growing tendency with increasing applied indentation load or relative indentation depth (see Fig. 6). The hardness of the coating is lower than the composite hardness in the composite region so it is confirmed that the system belongs „soft film-hard substrate“ composite system type.

Based on the results from: Fig. 4d, Table III (sample 3') and Fig. 6 conclude that a 30 % ultrasound amplitude (3.77 W/cm^2) is an optimal condition for copper ultrasonic deposition from electrolyte II.

D. Determination the copper coatings hardness and adhesion

Absolute hardness of the brass substrate, H_s, necessary to calculate the absolute hardness of the coating, is 1.41 GPa, according to our previous measurements [25].

The results of calculated film hardness according to the Chen–Gao model for the system ED Cu film on brass as the substrate were given in Table IV. Fitting of experimental datas (Fig. 7) were done in Matlab using the *cftool* command.

Films obtained with optimal condition of ultrasonic mixing appear harder than others deposited (sample No. 5). Copper coatings deposited from electrolyte with additives show lower hardness value then same coatings from electrolyte I.

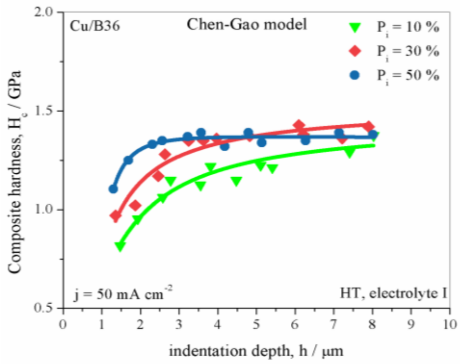


Fig. 7. The dependencies of the composite hardness of the Cu coatings, H_c on indentation depth, h calculated by Eq. (7) for various ultrasound mixing amplitude of electrolyte I.

TABLE IV
THE RESULTS OF CALCULATED FILM HARDNESS ACCORDING CHEN–GAO MODEL AND FITTING PARAMETERS (*A*, *B*, *C*) WITH ROOT MEAN SQUARE ERROR (*RMSE*)

No.	H_{coat}	*A*	*B*	*C*	*RMSE*
1	1.381	1.446	-7.20	-55.2	0.005
3	1.399	1.527	-5.46	-12.3	0.050
5	1.466	1.494	-1.52	-54.1	0.021
3'	1.357	1.412	-4.13	-43.0	0.038

The increase in hardness of electrodeposits obtained with ultrasonic agitation has been associated with the production of deposits with a fine grain size and higher grain-packing density [24]. Equation (9) was used to calculate the critical reduced depth b for the system of thin ED Cu films on the brass as the substrate. The value of the parameter n in the CG model can be 1.2 or 1.8, for soft films 1.8 is taken [12, 17-19].

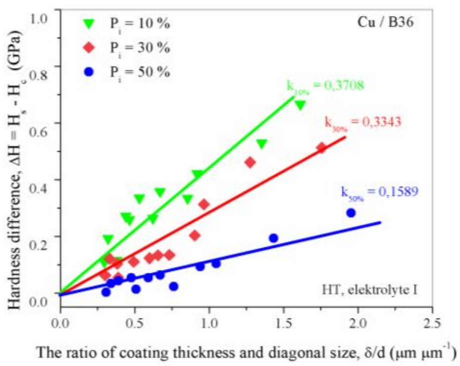


Fig. 8. Hardness difference vs. ratio between the film thickness and the indentation diagonal for copper films on the brass as a substrate. The slope value, k, ED Cu films (electrolyte I) for different ultrasonic mixing amplitude % is shown.

According to the model Chen–Gao (C–G) in Fig. 7 and Fig. 8 the values of slope, k (its value is used to calculate the critical reduction depth, b) for ED Cu films on brass deposited with different ultrasonic amplitude (10, 30 and 50 %) were shown. It was noted higher value of the slope, k, for the sample deposited at lower ultrasound intensity values.

The slopes values, k, (from Fig.7) and the absolute hardness values of the film (from Table IV) were used for the calculation the critical reduction depth, *b*, the calculated values are shown in the Table V.

TABLE V
THE RESULTS OF CALCULATED CRITICAL REDUCTION DEPTH, ACCORDING TO CHEN–GAO MODEL

No.	electrolyte	*k*	*b*
1	I	0.3708	2.995
3	I	0.3343	2.736
5	I	0.1589	1.165
3'	II	0.2872	4.777

Comparison of the adhesive strength by the adhesion parameter, *b*, can be noticed that higher value corresponds to the ED Cu films from electrolyte II than ED Cu films from electrolyte I. That means that better adhesion properties have ED Cu film with additives then ED Cu films without additives. When additives are added, more intense ultrasonic mixing power contributes to an increase in the adhesive strength of the substrate film and a better dissolution of the additives in the basic electrolyte. In the case of no additives, the adhesion value is almost unchanged with the variation of the mixing power.

V. CONCLUSION

Based on the microhardness measurements, it was shown that the same film with different structural properties giving different mechanical response depending on the electrochemical parameters during synthesis. The copper films deposited on brass from basic electrolyte have higher composite and film hardness values then sonoelectrodeposited copper films from electrolyte with additives for same deposition parameters and projected thickness.With increasing value of applied amplitude ultrasound mixing of electrolyte I, composite hardness value increasing, too. Adhesion properties is better for copper films from electrolyte II on brass then copper films from electrolyte I, because the values of *b* are larger for system ED Cu (electrolyte II)-brass then ED Cu (electrolyte I-brass). Better adhesion properties shown copper film deposited with a higher mixing condition.

ACKNOWLEDGMENT

This work was funded by Ministry of Education, Science and Technological Development of Republic of Serbia, Grant No.451-03-9/2021-14/200026.

REFERENCES

[1] L. C. Hagenson, L. K. Doraiswamy, "Comparison of the effects of ultrasound and mechanical agitation on a reacting solid-liquid system," *Chem. Eng. Sci.*, vol. *53*, no. *1*, pp. *131–148*, Apr., 1998.

[2] L. Fathyunes, J. Khalil-Allafi, "Effect of employing ultrasonic waves during pulse electrochemical deposition on the characteristics and biocompatibility of calcium phosphate coatings," *Ultrason. Sonochem.*, vol. *42*, pp. *293–302*, Apr., 2018.

[3] J. M. Costa, A. F. A. Neto, "Ultrasound-assisted electrodeposition and synthesis of alloys and composite materials: A review," *Ultrason. Sonochem.*, vol. *68*, pp. *105193*, Nov., 2020.

[4] C. T. Walker, R. Walker, "Effect of ultrasonic agitation on some properties of electrodeposits," *Electrodeposition Surf. Treat.*, vol. *1*, no. *6*, pp. *457–469*, Jul., 1973.

[5] A. K. Behera, A. Mallik, "Ultrasound assisted electroplating of nano-composite thin film of Cu matrix with electrochemically in-house synthesized few layer graphene nano-sheets as reinforcement," *J. Alloys Compd.*, vol. *750*,, pp. *587–98*, Mar., 2018.

[6] Y. Niu, J. Wei, Y. Yang, J. Hu, Z. Yu, "Influence of microstructure on the wear mechanism of multilayered Ni coating deposited by ultrasound-assisted electrodeposition," *Surf. Coat. Technol.*, vol. *210*, pp. *21–27*, Oct., 2012.

[7] A. Mallik, B. C. Ray, "Residual stress and nanomechanical properties of sonoelectrodeposited Cu films," *Surf. Eng.*, vol. *27*, no. *7*, pp. *551–556*, Aug., 2011.

[8] I. Tudela, Y. Zhang, M. Pal, I. Kerr, A. J. Cobley, "Ultrasound-assisted electrodeposition of composite coatings with particles," *Surf. Coat. Technol.*, vol. *259*, no. *C*, pp. *363–373*, Nov., 2014.

[9] K. Hedayati, G. Nabiyouni, "Surface roughness analysis and magnetic property studies of nickel thin films electrodeposited onto rotating disc electrodes," *Appl. Phys. A*, vol. *116*, pp. *1605–1612*, Feb., 2014.

[10] I. O. Mladenović, J. S. Lamovec, V. B. Jović, M. M. Obradov, D. G. Vasiljević-Radović, N. D. Nikolić, V. J. Radojević, "Mechanical characterization of copper coatings electrodeposited onto different substrates with and without ultrasound assistance," *J. Serb. Chem. Soc.*, vol. *84*, no. *7*, pp. *729–741*, Mar., 2019.

[11] L. Bonou, M. Eyraud, R. Denoyel, Y. Massiani, "Influence of additives on Cu electrodeposition mechanisms in acid solution: direct current study supported by non-electrochemical measurements", *Electrochim. Acta*, vol. *47*, no. *26*, pp. *4139–4148*, Oct., 2002.

[12] L. Magagnin, R. Maboudian, C. Carraro, "Adhesion evaluation of immersion plating copper films on silicon by microindentation measurements," *Thin Solid Films*, vol. *434*, no. *1*, pp. *100–105*, Jun., 2003.

[13] J. Lamovec, V. Jović, D. Ranđelović, R. Aleksić, V. Radojević, "Analysis of the composite and film hardness of electrodeposited nickel coatings on different substrates," *Thin Solid Films*, vol. *516*, no. *23*, pp. *8646–8654*, Oct., 2008.

[14] J. Lesage, D. Chicot, A. Pertuz, P. -Y. Jouan, N. Horny, A. Soom, "A model for hardness determination of thin coatings from standard micro-indentation tests,", *Surf. Coat. Technol.*, vol. *200*, no.*1–4* , pp. *886–889*, Oct., 2005.

[15] D. Chicot, J. Lesage, "Absolute hardness of films and coatings", *Thin Solid Films*, vol. *254*, no. *1–2*, pp. *123-130*, Jan., 1995.

[16] J. Lesage, A. Pertuz, D. Chicot, "A New Method to Determine the Hardness of Thin Films," *Revista Matéria*, vol. *9*, no. *1*, pp. *12–19*, Jan. 2004.

[17] M. Chen, J. Gao, "The adhesion of copper films coated on silicon and glass substrates," *Mod. Phys. Lett. B*, vol. *14*, no. *3*, pp. *103–108*, Mar., 2000.

[18] J. L. He, W. Z. Li, H. D. Li, "Hardness measurement of thin films: Separation from composite hardness," *Appl. Phys. Lett.* vol. *69*, no. *10*, pp. *1402–1404*, Jun., 1996.

[19] Q. Hou, J. Gao, S. Li, "Adhesion and its influence on micro-hardness of DLC and SiC films," *Eur. Phys. J. B*, vol. *8*, pp. *493–496*, Apr., 1999.

[20] N. D. Nikolić, Z. Rakočević, K. I. Popov, "Reflection and structural analyses of mirror-bright metal coatings," *J Solid State Electrochem.*, vol. *8*, pp. *526–531*, Dec., 2004.

[21] http://gwyddion.net/

[22] Y. Li, J. Yang, Z. Pan, W. Tong, "Nanoscale pore structure and mechanical property analysis of coal: An insight combining AFM and SEM images," *Fuel*, vol. *260*, pp. *116352*, Jan., 2020.

[23] *Metallic Materials—Vickers Hardness Test—Part 1: Test Method*, ISO 6507-1-2005, International Organization for Standardization: Geneva, Swizterland, 2005.

[24] P. M. Ashokkumar, "*Theoretical and Experimental Sonochemistry Involving Inorganic Systems*", *2011*th ed. New York, US: Springer, 2014.

[25] I. O. Mladenović, N. D. Nikolić, J. S. Lamovec, D. G. Vasiljević-Radović, V. J. Radojević, "Application of the Composite Hardness Models in the Analysis of Mechanical Characteristics of Electrolytically Deposited Copper Coatings: The Effect of the Type of Substrate," *Metals*, vol. *11*, no. *1*, pp. *111*, Jan., 2021.

Magnetic Field Generator For Simulation of a Vehicle Movement For a Wide Range of Velocities

Milan Stojanović, *Student Member, IEEE*, Jana Vračar, *Student Member, IEEE*, Ilija Neden Dimitriu, Ljubomir Vračar

***Abstract* — This paper describes the LTSpice model of a magnetic field generator needed for testing a vehicle detector. This simulator reproduces magnetic field distortion equal to a vehicle when passing over the detector located in the road in real conditions. The electric circuit of the simulator described in this paper is a solenoid, PWM generator, and filter capacitors. The different values of capacitors are given for different vehicle velocities. The switching matrix is used for selecting the appropriate capacitor values to achieve simulation of different vehicle velocities.**

***Index Terms* — Magnetic field generator, magnetic field distortion, LTSpice model, vehicle detection.**

I. Introduction

Vehicle detection is an essential part of traffic systems, especially in places of high population density. This system is used to get data about available parking spots and analyze traffic flow [1, 2]. Vehicle detection can be done by analyzing the magnetic field changes, which occur due to the vehicle presence [3]. Fig. 1 shows the placement of the magnetic field sensor on the road, and the changes of the magnetic field due to the passing vehicle over the sensor.

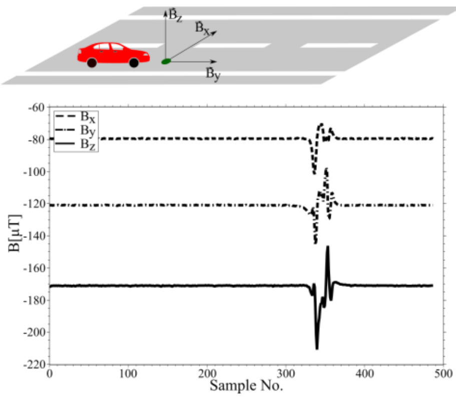


Fig. 1. The placement of the magnetic field sensor on the road, and changes of the magnetic field induced by the vehicle.

Milan Stojanović is with the Department of Microelectronics, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia (e-mail: milan.stojanovic@elfak.ni.ac.rs).
Jana Vračar is with the Department of Microelectronics, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia (e-mail: jana.vracar@elfak.ni.ac.rs).
Ljubomir Vračar is with the Department of Microelectronics, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia (e-mail: ljubomir.vracar@elfak.ni.ac.rs).
Ilija Neden Dimitriu is with the Public Enterprise "Roads of Serbia", 11000 Belgrade, Serbia (e-mail: ilija.nedendimitriu@putevi-srbije.rs).

As shown, the vehicle induces magnetic field distortion along all axes (*x*, *y*, *z*). Determining the vehicle presence can be done by analyzing the measured values from one [4] or several axes [5]. The measured results show that when the vehicle passes over the sensor, it creates bigger changes in the magnetic field along the *z*-axis. Therefore, the values measured along this axis are taken within the analysis usually.

The reliability of the detector needs to be tested before installing it on the road. That includes testing the success of the detection of different types of vehicles moving at different velocities. A simulator capable of reproducing a magnetic change identical to vehicle-induced is useful to increase the efficiency of the testing process. The solenoid is the part of the simulator that generates the magnetic field, and this field depends on the waveform of the current flow through the solenoid, and it is directly proportional to it. The current of an arbitrary waveform can be obtained using the PWM generator of the variable duty cycle (DTC) [6, 7]. If that generator drives the solenoid, the magnetic field of arbitrary shape is created. The rest of the paper shows the magnetic field simulator, described by electrical circuits in the software tool LTSpice. The current flow through the solenoid changes as the magnetic field due to the passing of the vehicle.

II. The Theoretical Principle of Simulator Operation

Fig. 2 shows a relation between the current changes through the solenoid ΔI, and the DTC of PWM voltage, at the end of the pulse. As can be seen, the increase of DTC contributes to an increased value of ΔI. The showed values of ΔI are absolute, so the value of the current flow is equal to zero before a pulse arrival.

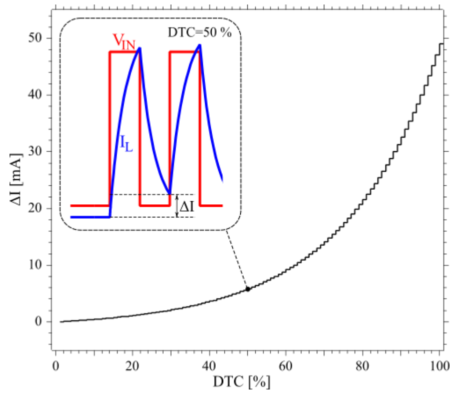


Fig. 2. Influence of duty cycle of PWM voltage to the intensity of the current through the solenoid.

The current flow through the solenoid is changed, by changing the duty cycle of PWM voltage. The current intensity can be controlled in this way, but the current always flows in the same direction. The coil should be connected to the two PWM generators to allow the change in the direction of the current flow and thus generate changes in the magnetic field as the vehicle does. Fig. 3 shows a simplified schematic of the simulator.

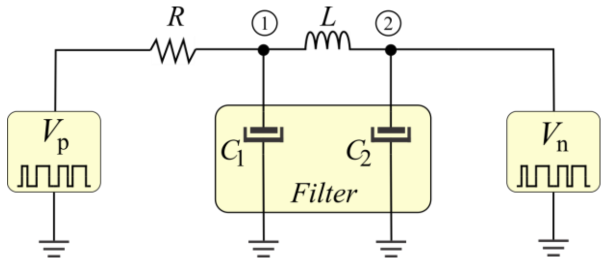


Fig. 3. Simplified schematic of the magnetic field simulator.

The simulations of current flow were done using LTSpice software. The ideal solenoid inductivity is 100 mH, and that value is used in the software. The resistor R limits the current intensity (magnetic field), and the value is 100 Ω. The capacitors C_1 and C_2 form a filter that suppresses sudden changes in a current flow when voltage level changes of both generators.

PWM generators are realized as PWL (Piecewise Linear) generators in LTSpice. The parameters of PWL generators were generated inside textual files. The textual file of every PWL generator contains couples of points (x, y), where x represents the time, and y represents voltage values. The generator makes pulses by reading values from the file.

A. PWM generators

The magnetic field that presents the passing of one vehicle is created by driving the solenoid by 100 pulses from the generators. The V_p makes positive, and V_n negative changes to the magnetic field.

The duty cycle of the PWM voltage of both generators is proportional to the changes in the magnetic field. The frequency of PWM voltage depends on the duration of the magnetic field distortion.

The duration of one pulse is calculated as

$$t_{pulse} = \frac{t}{100},$$

where t represents the time of magnetic field distortion. In this case, the distortion is induced by a vehicle, so t can be replaced by $t_{vehicle}$ and this value depends on the length ($l_{vehicle}$) of the vehicle and its velocity ($v_{vehicle}$)

$$t = t_{vehicle} = \frac{l_{vehicle}}{v_{vehicle}}.$$

So the duration of one pulse is

$$t_{pulse} = \frac{d_{vehicle}}{100 \cdot v_{vehicle}},$$

and the frequency of the pulse signal is equal to

$$f_{pulse} = \frac{1}{t_{pulse}} = 100 \cdot \frac{v_{vehicle}}{d_{vehicle}}.$$

The length of the vehicle is considered to be 4 m. The textual files of both PWL generator V_p and V_n are created based on the value of t_{pulse}. One pulse is defined by four points (x, y), as shown in Fig. 4.

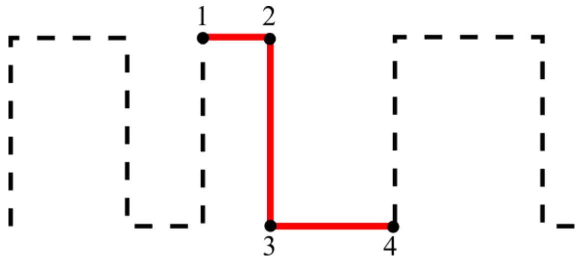


Fig. 4. Points that determine the frequency and duty cycle of one pulse of PWM voltage.

The x_1, y_1 are coordinates of the first point, x_2, y_2 the second, et cetera. The values of coordinates of all four points are calculated:

$$\begin{gathered}
(x_1, y_1) = \left((i-1)\cdot t_{pulse} + \Delta t, V_{ON}\right) \\
(x_2, y_2) = \left((i-1)\cdot t_{pulse} + DTC_j[i]\cdot\frac{t_{pulse}}{100} + \Delta t, V_{ON}\right) \\
(x_3, y_3) = \left((i-1)\cdot t_{pulse} + DTC_j[i]\cdot\frac{t_{pulse}}{100} + 2\Delta t, 0\right) \\
(x_4, y_4) = \left(i\cdot t_{pulse}, 0\right) \\
j \in \{p, n\}, i \in \{1, N\},
\end{gathered}$$

where N represents the number of pulses. Δt is the time interval needed to change the value of the generator voltage from 0 to V_{ON} and vice versa. This interval value is much smaller than the value of the t_{pulse}, so it does not affect the accuracy of the frequency of the generated voltage. The noted equations apply if the condition is met

$$DTC_j[i] \neq 0.$$

Otherwise, the voltage value at points 1 and 2 is equal to zero, and only point 4 within the current pulse is generated.

B. Filter capacitors

The capacitance values of C_1 and C_2 are equal, because the frequency of both generators is equal, and they are determined experimentally. Fig. 5 shows the current waveform through the solenoid for vehicle density of $1\frac{m}{s}$ and three different values of capacitors C_1 and C_2.

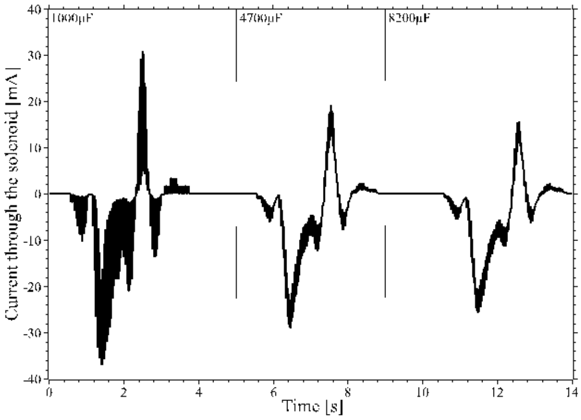


Fig 5. The waveforms of the current through the solenoid for vehicle velocity of $1\frac{m}{s}$ and three different values of C_1 and C_2.

If the capacitors C_1 and C_2 have a capacitance of 1000 μF there are unwanted high-frequency changes in the waveform of the current through the solenoid. These changes should be "smoothed" to get the required shape of the generated magnetic field.
In the other case, if the capacitors have too large values, such as 8200 μF, the amplitude of the current waveform becomes attenuated. So, the simulation shows that the appropriate values of C_1 and C_2 are 4700 μF for shown vehicle velocity.

In Fig. 6. the current through the solenoid for vehicle velocity from $1\frac{m}{s}$ to $5\frac{m}{s}$ is shown.

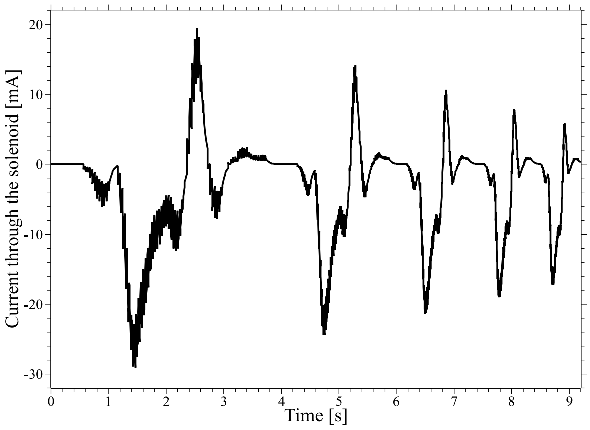


Fig. 6. The waveforms of the current through the solenoid for vehicle velocity from $1\frac{m}{s}$ to $5\frac{m}{s}$. The value of vehicle velocity increases from the left to the right side.

There are two problems when a vehicle's velocity rises. The amplitude of the current waveform decreases, but the waveform shape is degraded also. That happens because the capacitance values of C_1 and C_2 are too large. It is required to determine appropriate values of capacitors for all vehicle velocities, or for ranges of the velocity that is supposed to be simulated.

That is done experimentally, by creating the PWM generators of various frequencies. Table I shows the values of capacitors C_1 and C_2 which are depending on the vehicle velocity, whose movement is simulated.

TABLE I
VALUE OF FILTER'S CAPACITORS FOR DIFFERENT VEHICLE VELOCITY

Vehicle velocity [m/s]	Capacitance C_1 and C_2 [uF]
1	4700
2	1800
3	820
4 - 5	470
6 - 7	220
8 - 10	150
11 - 15	120
16 - 20	68
20 - 30	47
30 - 50	22

Capacitor values should be selected before starting the simulation.

III. SWITCHING MATRIX AND RESULTS

Changing capacitor values without interrupting the simulation and manual reconnection of new components can be made using the switching matrices, shown in Fig. 7. In this way, the capacitor value can be selected depending on the frequency of the PWM signal. It provides the ability to simulate an unlimited number of vehicles of different lengths and velocities.

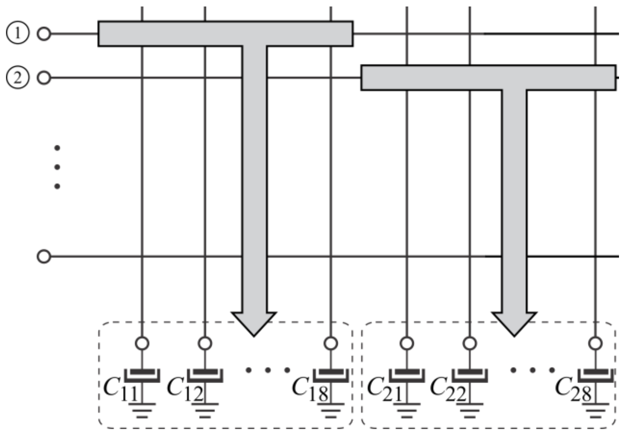


Fig. 7. The simplified block schematic of the switching matrix. The first and the second index of capacitors denotes the row and column of the switching element, respectively.

The complete switching matrix contains eight rows and 16 columns. In this case, two rows and the 8 columns of every row are used.

The necessary row number of the matrix determines the number of different capacitances connected simultaneously. Table I shows that the number of different values of capacitor capacitance is ten. The matrix contains eight columns, allowing the connection of up to eight various capacitors. The first row of the matrix is used instead of the capacitor C_1, and the second matrix instead of C_2. The capacitor values from C_{11} to C_{18} are shown in Table I, there the C_{11} is the biggest value. The capacitors connected in both rows are equal. The capacitance of C_{11} is equal to C_{21}, the C_{12} is equal to C_{22}, and so on. Because there are eight different capacitors in the circuit, and there are ten values that are needed for shown velocity range, two capacitances are missing. In this case, the capacitors of 68 µF and 220 µF are not connected because these values can be obtained as a combination of others. Two remaining capacitances are obtained by a parallel connection of connected capacitors, 68 µF as the sum of 47 µF and 22 µF, and 220 µF as the sum of 150 µF, 47 µF, and 22 µF.

The selection of the capacitors that need to be connected to the output is done by the components inside the matrix. In Fig. 8 is shown the simplified schematic of one row of the switching matrix. The microcontroller controls all rows and is positioned left out of the schematic for better visibility.

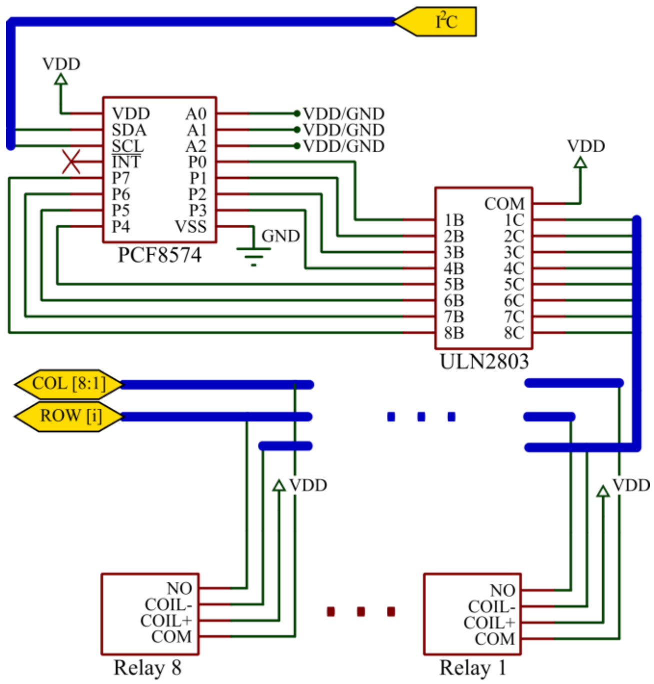


Fig. 8. The simplified schematic of one row of the switching matrix.

The input stage of a row is the I/O expander [8] that receives data from the microcontroller via I^2C communication. The received one byte of data determines which switching elements will be activated. This information is forwarded to the driver [9], which activates one or more appropriate switching elements (relays) [10]. The capacitors are connected to the output of the matrix, after the activation of relays. The PWM generators, after the selection of the capacitors, are activated. If the next simulated vehicle has velocity from another range, the appropriate capacitance is set again, and the simulation is repeated.

The current waveform through the solenoid is shown in Fig. 9, wherein the corresponding capacitor is used for each simulated vehicle velocity.

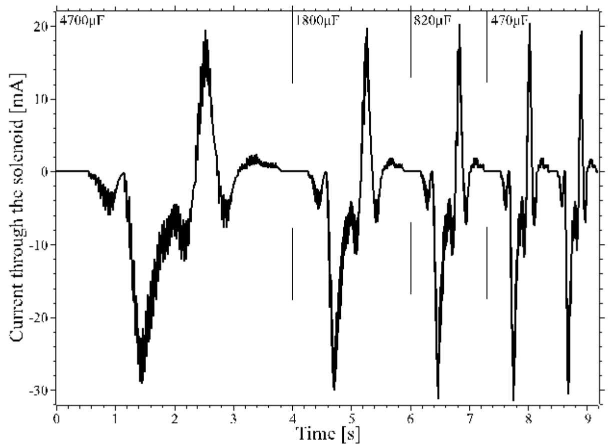


Fig. 9. The current through the solenoid for vehicle velocity from 1 m/s to 5 m/s, with a corresponding capacitor used for each simulated vehicle velocity.

Comparing Fig. 9 and Fig. 6 can be seen that using of switching matrix contributes to getting the current waveforms without degradations. The current through the solenoid has the same waveform for all simulated vehicle velocities. The current amplitude is also unchanged and does not depend on a vehicle's velocity.

IV. Conclusion

In this paper simulation for replicating a magnetic change identical to vehicle-induced is described. The simulator is described by the electrical circuit in LTSpice. The part of the simulator is a solenoid that generates the magnetic field. The PWM generators of variable DTC drive the solenoid to generate the current flow of arbitrary shape. That current flows through the solenoid and the magnetic field of the same waveform is generated.

Connecting the solenoid to the two PWM generators results in a change in the direction of the current flow. That change generates changes in the magnetic field equal to vehicle-induced changes in the magnetic field.

Degradation of generated magnetic field for different velocities are regulated by capacitor filter, with various values of capacitors C_1 and C_2. The values for capacitors are determined experimentally, and a part of the switching matrix for selecting the values for capacitors is used. When using the appropriate capacitor value, the amplitude and waveform of the current through the solenoid are unchanged for a wide range of vehicle velocities.

The future research will be based on expanding the scope of simulations for different types of vehicles. Also, one of the next steps in researching is the practical realization of the described simulator.

ACKNOWLEDGMENT

This paper is supported by the Serbian Ministry of Education, Science and Technological Development No. 451-03-9/2021-14/200102.

The described research is done in cooperation with Public Enterprise "Roads of Serbia", Serbia.

REFERENCES

[1] J. Gu, Z. Zhang, F. Yu and Q. Liu, "Design and implementation of a street parking system using wireless sensor networks," IEEE 10th International Conference on Industrial Informatics (INDIN), Beijing, China, pp. 1212-1217, 25-27 July 2012.

[2] E. Sifuentes, O. Casas, and R. Pallas-Areny, “Wireless Magnetic Sensor Node for Vehicle Detection With Optical Wake-Up”, *IEEE SENSORS JOURNAL*, vol. 11, no. 8, pp. 1669-1676, August 2011.

[3] Michael J. Caruso, Lucky S. Withanawasam, “Vehicle Detection and Compass Applications using AMR Magnetic Sensors”, Honeywell, SSEC, 12001 State Highway 55, Plymouth, MN USA 55441

[4] S. Y. Cheung, S. Coleri, B. Dundar, S. Ganesh, C-W Tan and P Varaiya, “Traffic Measurement and Vehicle Classification with Single Magnetic Sensor”, *Transportation Research Record*, vol. 1917, no. 1, pp.173-181, 2005.

[5] Q. Wang, J. Zheng, H. Xu, B. Xu and R. Chen, "Roadside Magnetic Sensor System for Vehicle Detection in Urban Environments," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 5, pp. 1365-1374, May 2018.

[6] Q. Zhong, B. Zhang, H-Y Yang, J-E Ma, R-F Fung. “Performance analysis of a high-speed on/off valve based on an intelligent pulse-width modulation control”, *Advances in Mechanical Engineering*, vol 9, no. 11, November 2017.

[7] V. Steinsland, “Modeling and Control of Retarder using On/Off Solenoid Valves”, Masters' Degree Project, Stockholm, Sweden April 2008.

[8] PCF8574 Remote 8-Bit I/O Expander for I^2C Bus, Datasheet, https://www.ti.com/lit/ds/symlink/pcf8574.pdf?ts=1621942328093&ref_url=https%253A%252F%252Fwww.google.com%252F

[9] ULN2803A Darlington Transistor Arrays, Datasheet, https://www.ti.com/lit/ds/symlink/uln2803a.pdf?ts=1621921786139&ref_url=https%253A%252F%252Fwww.startpage.com%252F

[10] HE3600 Miniature Single In-line Reed Relay, Datasheet https://eu.mouser.com/datasheet/2/240/Littelfuse_Reed_Relays_HE3600_Datasheet_pdf-876890.pdf

Active Matrix Liquid Crystal Display – AMLCD Switching Time Measurements

Branko Livada, *Member SPIE*

***Abstract*—Liquid crystal displays (LCD) are currently dominant in visual content presentation applications. In the case of video content presentation LCD switching time determines image appearance and perception. The photometric and temporal properties of LCDs are unfamiliar to many practitioners, especially switching time temperature dependence. In specific applications as ruggedized military displays temperature dependence is critical parameter for operation on low temperatures. Current liquid display technology does not provide proper operation at low temperatures so additional heater should be incorporated into display design. The temperature dependence of switching time provides initial data determining heater design parameters. The short review of the LCD switching time theory is presented. The AMLCD switching time over temperature, method is described. The selected measurement results are presented and discussed.**

***Index Terms*—liquid crystal display; active matrix liquid crystal display; switching time; rise time; fall time; measurement method; temperature dependence.**

I. Introduction

Twenty first century is century of liquid crystal (LC) displays (LCDs) that have been widely applied from small-size mobile devices to large-size televisions. Various methods have been developed for higher and advanced performance of active-matrix LCDs (AMLCDs) to provide superior application results [1]-[7]. Some of the most important achievements are: fast response LC providing high frame rate driving, backlight dimming technology for low consumption and high contrast ratio, wide viewing angle LC modes, high temperature range of operation etc. The majority of advancements have been achieved through advancements in LC science and technology [8]-[10].

The deep knowledge about LC [11]-[13] physical, optical and electro-optical properties is extensively used in design and operation of LCD pixel. One of the critical parameters is LCD pixel switching performances influencing proper video content reproduction, from the well–known issue of motion blur, to photometric settings and calibrations. Displays with faster transitions (switching) between different luminance levels are considered as supperior in the case of video content presentation. Because of that LC optical switching performances are studied theoretically and related measurement techniques have been developed. The basic data about LCD switching time is summarized in separate chapter.

Branko Livada is with the Vlatacom Institute of High Technologies, 5 Bulevar Militina Milankovića, 11070 Belgrade, Serbia (e-mail: branko.livada@ vlatacom.com).

In the specific application as ruggedized military displays [14]-[18] some of performances are critically important for application. Among the other parameters switching time is slightly more important, because it has influence on display design in addition to operation parameters. Current LC technology do not provide proper LCD operation at low temperatures as required in military applications, so additional heater should be built in display to provide proper display operation at low temperatures. The temperature dependence of the switching time determines the heater design and consumption sufficient to provide display proper operation at low temperatures.

In this paper measurement method applicable for display switching time measurements over display operation temperature range is described. The selected measurement results illustrating the switching time temperature dependence for several AMLCD panels are presented. Also, these results provide initial data for LCD heater design.

II. AMLCD Switching time

Liquid crystal occupies a small portion in an LC device, but plays a key role in determining the device performances. The LC material and molecular alignment jointly determine the device contrast ratio, operation voltage, response time, viewing angle, and operating temperature.

The LCD response time consideration has two key aspects:

(a) How to design liquid crystal composition to have desired response – switching time values? This is not topic of this article, but some aspects will be mentioned just for better understanding. More details could be find in references [19]-[29] dealing with theoretical studies and modeling of the switching time behavior in different LC materials and LCD architectures.

(b) What is the influence of the switching time on presented video content appearance and perception? This is important to define limiting values depending on application [30], [31].

The LCD response time depends on different factors as illustrated in TABLE I. All off listed parameters depends on temperature, so LC scientists have a hard task to design LC composition having desirable temperature dependence on temperature.

The temperature dependence of dielectric anisotropy - $\Delta\varepsilon$ is limiting factor at high temperatures. Usually as temperature increases, $\varepsilon_{//}$ decreases while $\varepsilon_{\perp}$ increases gradually, resulting in a decreased $\Delta\varepsilon$. As $T > T_c$, the isotropic state is reached and the dielectric anisotropy is vanished, with no birefringence. Critical temperature T_c is high temperature at which LC lost its birefringence properties and LCD stop to work. Typical values for T_c are in the range 70°C up to 100°C.

TABLE I
THE EFFECTS OF DIFFERENT FACTORS ON RESPONSE TIME

Factors	T_{on}	T_{off}
Viscosity (γ1) ↓	↓	↓
Elastic constants (Kii) ↑	↑	↓
Dielectric anisotropy (Δε)↑	↓	↓
Thickness (d) ↓	↓	↓
Pretilt angle (θ0) ↓	↑	↓
Anchoring energy (W) ↑	↑	↓
Temperature (T) ↑	↓	↓
Voltage (V) ↑	↓	↓

The viscosity and elastic constants are limiting factors at low temperatures causing slower motion of LC molecules when electrical field is applied to change the order of LC in device.

The LC electro optical properties (time and temperature dependence) influence the display pixel response characteristics defining display video content appearance. One of the most important effects is motion blur appearance.

The motion-induced distortion [32], [33] of a visual target moving across an electronic display screen perceived as blurring of initially sharp edges by a human observer (motion blur) is caused by two effects:

- characteristics of the electro-optical response of the display to changes in video signal
- integration properties of the of the human visual system while

The increased response times of LCD pixel corrupt the visual quality of moving objects and thus contribute to motion blur.

III. AMLCD SWITCHING TIME MEASUREMENT METHOD

Motion blur influences display applicability on the different way depending on presented video content properties. Because of that motion blur measurement methods [34], [35] should be designed on different ways as illustrated in Fig.1.

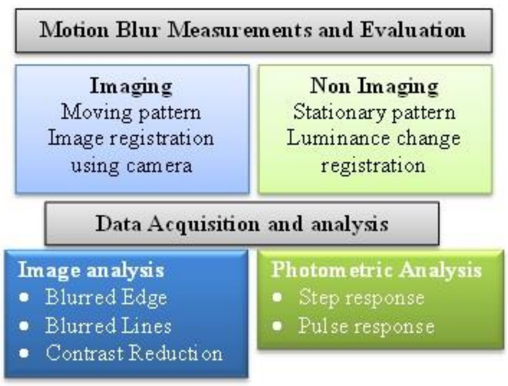


Fig. 1. Motion blur measurement and evaluation methods

Imaging methods are using selected moving patterns (blocks, lines, grille, grating etc.) presented on display and related image is recorded with tracking or stationary electronic cameras, and after that analyzed to quantify motion blur effects.

Non-imaging methods use stationary patterns on display surface but having temporary defined intensity change (pulse, step, periodic etc.) and pattern luminance level is measured using photometer [36].

In the case of ruggedized display presenting virtual instrument scale having moving bar edge proportional to the related quantity value it is important to have un-blurred moving edge to provide proper instrument reading. The LCD switching time value could be limitation factor. That type of display should be used reliably over all operation temperatures. As LC has known limitation for operation on low temperatures so LCD switching time temperature dependence should be known. This type of data is not presented in display data sheet so it is necessary that display designer provide that data.

Ideally, the measurement system should provide data about display panel brightness change versus temperature. To provide this it is necessary to avoid influence of the temperature to light sensor subsystem. For switching speed measurements, brightness change over time at set temperature should be recorded.

AMLCD Response Time is usually defined as sum of rise time (T_{on}) and fall time (T_{off}). T_{on} is defined as time necessary to change light signal from 10% value to 90% value when display pattern is changing from black to white. T_{off} is defined as time necessary to change light signal from 90% value to 10% value when display pattern is changing from white to black (See Fig.2).

We could only to compare response time (T_{on} & T_{off}) as defined previously. Also gray-to-gray level response time matters for video content presentation. In our case we are presenting moving bar instrument scale so it is good enough to measure only black to white level change.

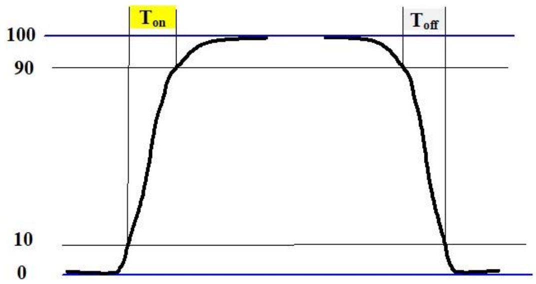


Fig. 2. T_{on} and T_{off} definition

Response Time is important parameter we need to know, to specify, at first place, heater requirements. Also, it is important parameter defining speed of response in the case of high frame rate imaging and fast motion artifacts in image. The way how it is defined does not allow us to have full sense what influence it has to image quality. Pixel response delay due to driving electronics is missing parameter.

Display temporal response was measured using a photodiode-based circuit and an oscilloscope. The photodiode was directed at switchable (black and white) square generated by our test program. A memory scope was used to record the photodiode response. Controlled temperature chamber is used to stabilize AMLCD panel temperature. Originally proposed measurement set-up is presented in Fig.3. It is proposed as best solution to use collection lens coupled with fiber-optical cable to provide that temperature have influence only on AMLCD panel while other parts of measurement chain are on room temperature. Following the facts that Si photodiode do not suffer of the response temperature dependence and that T_{on} , T_{off} data are extracted from relative photo signal values (see Fig.2) we selected simplified solution presented in Fig.4 and using collection lens and Si photodiode placed in temperature chamber and connected to photo current measurement circuit via coaxial cable.

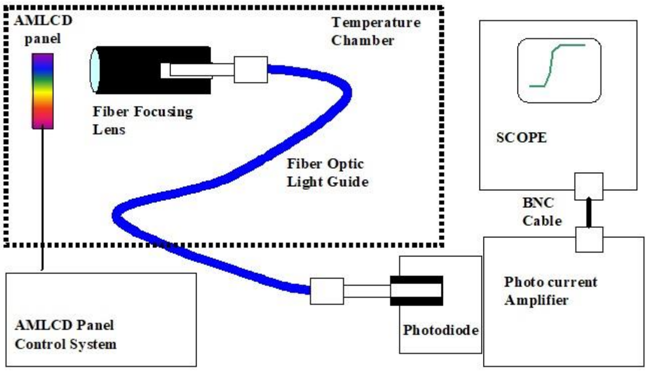


Fig. 3. Switching time measurement set-up

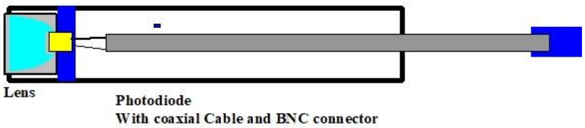


Fig. 4. Photodiode with collection lens

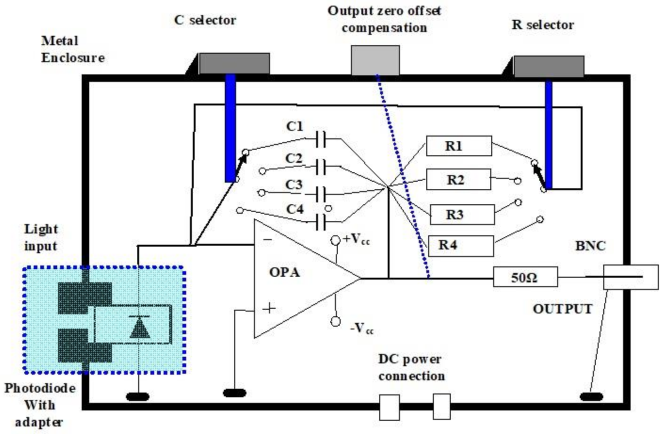


Fig. 5. Photocurrent to voltage electronic block schematic

Generated photocurrent is measured using photo current to voltage converter (trans-impedance operational amplifier) presented in Fig.4.

The light generated signal is generally noisy or involving high frequency components due to LCD backlight switching. It is good to filter the signal before recording to have more smooth line and easier determination of required 10% and 90% values. Because of that we use additional capacitance over load resistor for filtering, but RC constant should be at least ten times less than expected rise or fall time values.

The described photometric device and climatic chamber offers, flexible and easy evaluation of displays switching time in specific temperature environment between -40°C to +85°C.

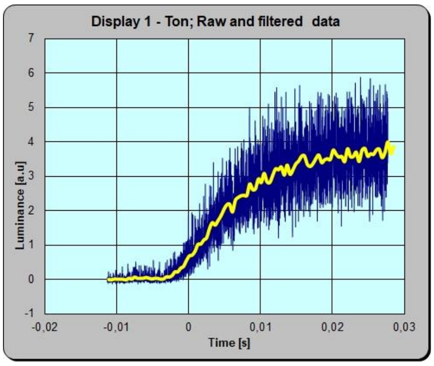
Sensor and amplifier were tested and calibrated using 586 fL uniform halogen lamp in integration sphere white source (color coordinates x=0,386, y=0,401). Calibration results are summarized and presented in TABLE II.

TABLE II
CALIBRATION RESULTS

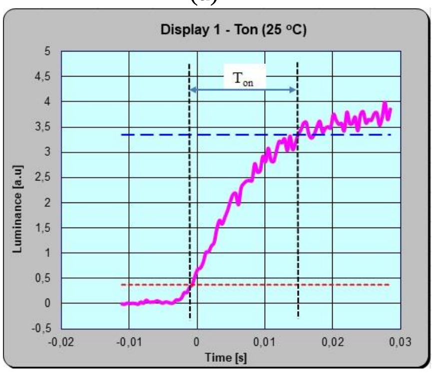
R [K]		Voltage Sensitivity [mV/fL]	Photo -Current Sensitivity [nA/fL]
R1	100	0.59	5.87
R2	388	2.28	5.88
R3	910	5.37	5.90
R4	3260	19.28	5.91
Average current sensitivity			**5.89**

IV. AMLCD SWITCHING TIME MEASUREMENT RESULTS

In the proposed paper, we present experimental results obtained on several commercial LCDs using a thermal enclosure allowing a required temperature range.



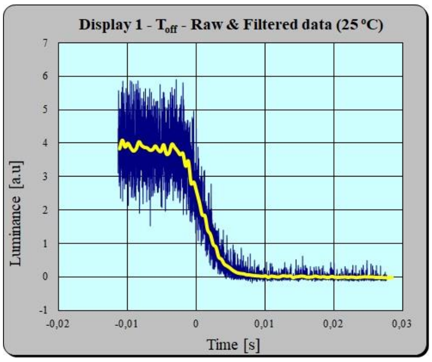
(a)



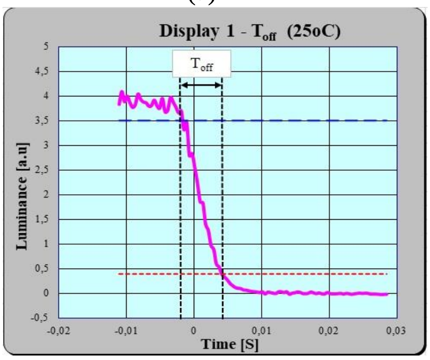
(b)

Fig. 6. T_{on} measurement signal (a) raw signal, (b) processed signal

The typical view of the recorded photometric signal during white to black and black to white switching is presented in Fig.6 and Fig.7.



(a)



(b)

Fig. 7. T_{off} measurement signal (a) raw signal, (b) processed signal

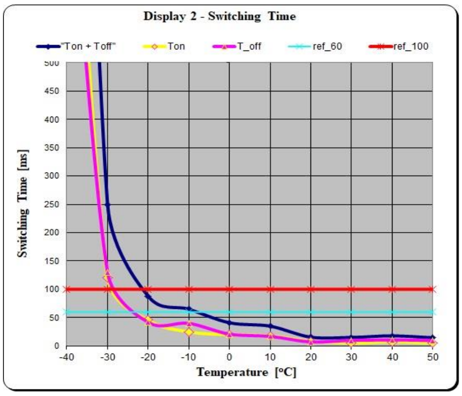


Fig. 8. Switching time vs. temperature measurement results (Display 2)

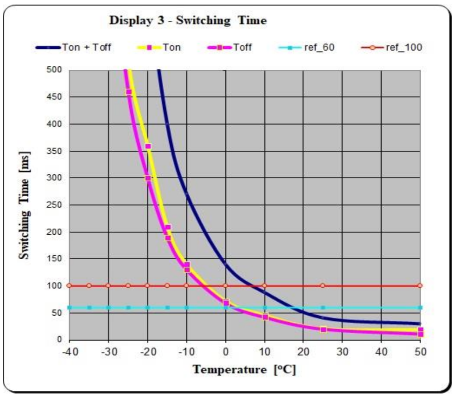


Fig. 9. Switching time vs. temperature measurement results (Display 3)

Measurement signal is exported as ASCI file and transferred to EXCEL spread sheet for further processing. Following recommendations from [37] and [38] moving averaging filter is successfully used for curve smoothing as illustrated in Fig.5 (a) and Fig.6(a).

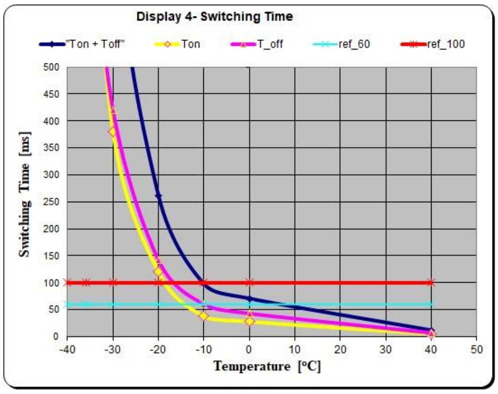


Fig. 10. Switching time vs. temperature measurement results (Display 4)

The summary of the measurement results for several AMLCD panels are presented in Fig.8 to Fig.10 together with two selected switching time critical values (60ms and 100ms). These results could be used to determine the lowest display operation temperature and provide starting data for heater design

V. Conclusion

AMLCD switching time depends on various LC parameters and LCD pixel design and operation. Also, switching time has influence on AMLCD panel operation. One of the most important switching time influences is motion blur appearance.

New light sensor aimed for response time measurements were built and tested and calibrated for photo current response.

Testing and calibration results show that it could be used successfully in all future response time measurements.

Switching time vs. temperature data are presented for several selected AMLCD panels and show that they have different critical temperature values. To use effectively these measurement results for heater design it is necessary to determine switching time (T_{on}, T_{off} or T_{on}+T_{off}) critical values to avoid motion blur influence according to display application.

It is shown that relatively simple and accurate method for switching time measurements could be designed and applied for aimed purpose when display critical switching time is known. The measurement methodology is successfully applied

In the case of display presenting virtual instrument scale having moving bar edge it was possible to determine critical switching time values. Using switching time vs. temperature measurement results heater design requirements could be derived.

References

[1] *D.-K.* Yang and S.-T. Wu, *Fundamentals of Liquid Crystal Devices* , John Wiley & Sons, Ltd., New York, USA, 2006

[2] Jiun-Haw Lee, David N. Liu and Shin-TsonWu, *Introduction to Flat Panel Displays*, JohnWiley & Sons Ltd., Chichester, UK, 2008

[3] Robert H. Chen, *Liquid Crystal Displays: Fundamental Physics and Technology,* by John Wiley & Sons, Inc., Hoboken, New Jersey, USA, 2011

[4] Deng-Ke Yang and Shin-Tson Wu, *Fundamentals of Liquid Crystal Devices*, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, England, UK, 2006

[5] Roll R.Hainich, Oliver Bimber, *DISPLAYS: Fundamentals and Applications*, 2nd edition, Taylor & Francis Group, LLC, CRC Press,Boca Raton, USA, 2017

[6] Willem den Boer, Active Matrix Liquid Crystal Displays: Fundamentals and Applications, Elsevier Inc., New York, USA, 2005

[7] Janglin Chen, Wayne Cranton, Mark Fihn (Eds.), *Handbook of Visual Display Technology*, Springer-Verlag, Berlin, Heidelberg, Germany, 2012

[8] Hirohisa Kawamoto, The History of Liquid-Crystal Displays, *Proc. IEEE*, Vol. 90, No. 4, pp.460-500, 2002

[9] Haiwei Chen & Shin-Tson Wu, Advanced liquid crystal displays with supreme image qualities, *Liquid Crystals Today*, 28:1, 4-11, 2019

[10] Askari Mohammad Bagher, Mirzaei Mahmoud Abadi Vahid, Mirhabibi Mohsen, A Review of Challenges in Display Technology, International *Journal of Electrical Components and Energy Conversion*, 3(2): 26-39, 2017

[11] D. Demus, J. Goodby, G.W. Gray, H.-W. Spiess, V Vill, *Physical Properties of Liquid Crystals*, Wiley-VCH, New York, Chichester, Brisbane; Singapore, Toronto, 1999

[12] Khoo, Iam-Choon: *Liquid crystals*, 2nd ed., John Wiley & Sons, Inc., Hoboken, New Jersey, USA, 2006

[13] Sergey V. Pasechnik, Vladimir G. Chigrinov, Dina V. Shmeliova, *Liquid Crystals: Viscous and Elastic Properties*, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2009

[14] Keith K. Niall (Editor), *Vision and Displays for Military and Security Applications: The Advanced Deployable Day/Night Simulation Project*, Springer Science+Business Media, LLC, Springer New York Dordrecht Heidelberg London, 2010

[15] Branko Livada, Radomir Jankovic, Nikolic, AFV Vetronics: Displays Design Criteria, *Strojniški vestnik - Journal of Mechanical Engineering (JME)*, vol.**58**, No6, 376-385, 2012

[16] Branko Livada, Ruggedized Avionic Displays Technology: Past, present, and future, *4th International Scientific Conference OTEH 2011*, Belgrade, 06-07, October 2011, 2011

[17] Branko Livada, "Avionic Displays", *Scientific Technical Review*, Vol. **62**, No.3-4,pp.70-79, 2012

[18] Branko Livada, Displays Metrology status and new challenges, *9th International Scientific Conference OTEH 2020*, Belgrade, 08-09, October 2020, 2020

[19] H Wohler, M. E. Becker, Numerical Modeling of LCD Electro-Optical performances, *Opto-Electronics Review*, 10(1), 23-33, 2002

[20] Fenglin Peng, Haiwei Chen, Fangwang Gou, Yun-Han Lee, Michael Wand, Ming-Chun Li, Seok-Lyul Lee, and Shin-Tson Wu, Analytical equation for the motion picture response time of display devices, *J. Appl. Phys*. 121, 023108 ,2017

[21] Lachezar Komitov, Liquid Crystal Polar State and Fast Switching, *IMID/IDMC/ASIA DISPLAY 2010 DIGEST, 18, 2010*

[22] Pierre Adam, Pascal Bertolino, Fritz Lebowsky, Mathematical modeling of the LCD response time, Journal of the Society for Information Display, 15 (8), pp.571-577, 2007

[23] Youngmin Cho, Chansoo Park, Jong-Man Kim, Seung-Hyuck Lee, and Seung-Woo Lee, A Behavioral Circuit Model of Active-Matrix Liquid Crystal Displays for Optical Response Simulation, *IEEE TRANSACTIONS ON ELECTRON DEVICES*, Vol. 59, No. 5, 1430-1438, 2012

[24] Jong-Man Kim, Seung-Hyuck Lee, and Seung-Woo Lee, Behavioral Model of Patterned Vertical Alignment Pixel in Active-Matrix Liquid Crystal Displays, *IEEE TRANSACTIONS ON ELECTRON DEVICES*, VOL. 61, NO. 11, 3783-3790, 2014

[25] S. H. Lee, S. L. Lee, and H. Y. Kim, Electro-optic characteristics and switching principle of a nematic liquid crystal cell controlled by fringe-field switching, *Applied Physics Letters* **73**, 2881-2883, 1998

[26] Pierre BOHER, Thierry LEROUX and Thibault BIGNON, Response Time and Viewing Angle Behavior of Liquid Crystal Displays versus Temperature, *Conference: Vehicle Displays*, Livonia, USA, September 2016

[27] Jong-Man Kim, Dong-Gun Lee, and Seung-Woo Lee, Temperature-Dependent Behavioral Model of Active-Matrix Liquid Crystal Displays for All Types of Liquid Crystals, *Journal of Display Technology*, Vol. 12, Issue 9, pp. 881-887, 2016

[28] Jong-Man Kim and Seung-Woo Lee, Temperature-dependent behavioral model of pixels in active-matrix liquid crystal displays with fringe-field switching mode, *Optical Engineering,* 54(9), 093104, 2015

[29] Jong-Man Kim and Seung-Woo Lee, Temperature-Dependent Behavioral Model of Twisted Nematic Pixel in Active-Matrix Liquid Crystal Displays, *SID 2015 DIGEST*, 1536-1539, 2015

[30] Tobias Elze, Thomas G. Tanner, Temporal Properties of Liquid Crystal Displays: Implications for Vision Science Experiments, *PLOS ONE*,| Volume 7, Issue 9, e44048, September 2012

[31] Jun Someya, Hiroaki Sugiura, Evaluation of liquid-crystal-display motion blur with moving-picture response time and human perception, *Journal of the SID* 15/1, 78-86, 2007

[32] Michael E. Becker, Motion-blur evaluation: A comparison of approaches, *Journal of the SID* 16/10, 989-1000, 2008

[33] Andrew B. Watson, Display motion blur: Comparison of measurement methods, Journal of the SID 18/2, 179-190, 2010

[34] Michael E. Becker, Motion Blur Measurement and Evaluation: From Theory to the Laboratory, SID07 DIGEST, 1122-1226, 2007

[35] Michael E. Becker, LCD Response Time Evaluation in the Presence of Backlight Modulations, *SID 08 DIGEST*, 1, 2008

[36] Michael E. Becker, Temporal characteristics of electronic display screens, DM&S 2013

[37] --, *Information Display Measurement standard*.- IDMS v.1.03, Society of Information Displays, June 01, 2012

[38] José Luis Guiñón, Emma Ortega, José García-Antón, Valentín Pérez-Herranz, Moving Average and Savitzki-Golay Smoothing Filters Using Mathcad, *International Conference on Engineering Education – ICEE 2007*, Coimbra, Portugal September 3 – 7, 2007

Hyper Focal Distance Application for Long Range Surveillance Camera Zoom Lens Focusing Settings

Saša Vujić, Member SPIE, Dragana Perić, *Member, IEEE*, Branko Livada, *Member, SPIE*

***Abstract*—Modern multi sensor surveillance systems integrate several multispectral imaging channels, laser range finder and positioning sensors (digital magnetic compass – north finding sensor and GPS sensor). Imaging channels use motorized zoom lenses providing convenient and fast field of view –FOV change. The fast FOV change to desired value keeping optimal image sharpness is provided through selected FOV pre-set settings using zoom lens position setting to desired values. In addition to FOV settings it is useful to pre-set lens focus motor position. The zoom lens hyper-focal distance determination for selected FOV and lens focusing motor position setting accordingly is used as pre-set lens parameter definition. The short review of the motorized zoom lens design and their basic properties is presented. The lens depth of field and hyper focal distance are discussed and basic formulas are derived. The zoom lens based imaging channel calibration procedures selection depends on application. We presented in detail hyper-focal distance based focusing motor parameter setting as one of the calibration procedures used in our multi sensor imaging systems.**

***Index Terms*—Surveillance systems, imaging system calibration, motorized zoom lens, depth of field, hyper focal distance**

I. Introduction

Multi sensor – multi spectral surveillance system users prefer application of the zoom lenses in the built in camera systems. Motorized zoom lens provides flexibility and controllability of the imaging conditions: changing the focal length, focusing distance, and aperture value to suit different fields of view (FOVs), depths of field (DOFs), and lighting conditions. Digital image sensor and application of the computer for surveillance system control provides technical means for convenient and accurate imaging data analysis and camera operation control using pre-set and calibration data generated in the laboratory.

Zoom lens calibration is described in open literature mainly for zoom lens application in photogrammetric measurement and machine vision application [1-8]. Camera calibration is a prerequisite for 3D imaging applications, providing data for 3D reconstruction. The calibration procedures could be identified as: standard calibration, self-calibration, photometric calibration and stereo-setup calibration. There are a lot of techniques applied for specific applications, but calibration of the surveillance zoom camera aimed for object pointing and tracking is poorly covered in the open literature. In this application [9] the most important is so called "zoom to FOV" calibration, involving lens optical axis stability calibration and line of sight – LoS control, using position control of the sighting reticle. Zoom lens focal length and image focus is set with digitally controlled positioning motors. "Zoom to FOV" calibration provides connection between motor position digital data and related zoom lens focal length or camera FOV. LoS stability control against zoom motor position is provided using data of optical axis displacement versus zoom motor position. This zoom lens calibration in surveillance systems provides accurate system aiming and target tracking. Sharp image is provided by additional camera focus control.

The long range surveillance system should provide wide field of regard – FOR (target search over wide area and different ranges) and possibility to provide high magnification in the same time. High magnification means that imagers use narrow field of view that is incompatible with wide FOR. This discrepancy could be overridden using zoom function. FOR search is done using wide FOV for target detection, and switch to narrow FOV for target recognition and identification. Human operator is usually involved in the system control chain. To support fast switching from wide FOV to suitable narrow FOV several zoom (FOV) pre-set positions are defined. To avoid additional time loss for focus adjustment it is convenient to pre-set focus motor position and provide sharp image immediately after switching to desired FOV. One of the possible solutions is to use idea about hyper-focal distance that is widely used in the cinema industry and photography [10].

In this article we are describing the basic design and properties of the motorized zoom lenses. We are pointing out zoom lens specifics leading to imperfections that should be corrected by calibration, together with definition of the main geometrical relations in the camera system. Also we described our simplified model for hyper-focal distance calculation applied for motorized zoom lens cameras used in our systems. The methodology for implementation of these results to camera pre-set parameters is explained.

Saša Vujić is with Vlatacom Institute of High Technologies, 5 Bulevar Milutina Milankovića, 11070 Belgrade, Serbia (e-mail: sasa.vujic@ vlatacom.com).

Dragana Perić is with Vlatacom Institute of High Technologies,, 5 Bulevar Milutina Milankovića, 11070 Belgrade, Serbia (e-mail: dragana.peric@ vlatacom.com).

Branko Livada is with Vlatacom Institute of High Technologies, 5 Bulevar Milutina Milankovića, 11070 Belgrade, Serbia (e-mail: branko.livada@ vlatacom.com).

In the case of zoom lens calibration and set-up, common camera calibration should be conducted first. Camera calibration is the process of determining the intrinsic parameters of the camera, including back focus distance.

A zoom lens contains moving lens groups providing focal length change that could be moved using linear or rotational stages powered by electrical motor. To provide high effectivity of zoom lens controllability, lens calibration and initial set-up should be performed.

The focus adjustments for selected zoom pre-set positions is the last step in camera setting.

II. Zoom Lens Design and Basic Properties

Imaging zoom lens is capable of producing, on a fixed plane (image plane), images whose magnification is continuously variable between two extreme values. The variation in magnification is achieved by changing the relative positions of the lens elements or lens groups built in the zoom system. The zoom lens design is complex [11-18], including complicated mechanical motion system design and optical design and optimization, depending on dedicated zoom lens system application. Because of that, zoom lenses were not considered for high quality imaging application until 1960, when started wider application in commercial photographing systems.

Two broad classifications of the zoom systems recognize the mechanically compensated type and the optically compensated type. In mechanical compensation design, the motions given to the elements are such as to ensure exact constancy of focal position, these movements being non-linear and achieved by complicated camera mechanisms. In an optically compensated design the motion of all moving elements is identical and linear, but desired focus can be achieved only at a finite number of zoom positions.

The development of the digital imaging sensors leading to the small sensor size accompanied with aspherical lenses provides leads to the much wider applications of the zoom lens systems [14]. Long range surveillance application in visible and infrared region, requiring continuous focus caused that mechanically compensated zoom lens using fine motorized controls dominates in practical applications.

The zoom lens generalized structure is presented in Fig.1 showing the key fixed and moving (controllable) components. The illustration of complexity of zoom lens design is shown in in Fig.2.

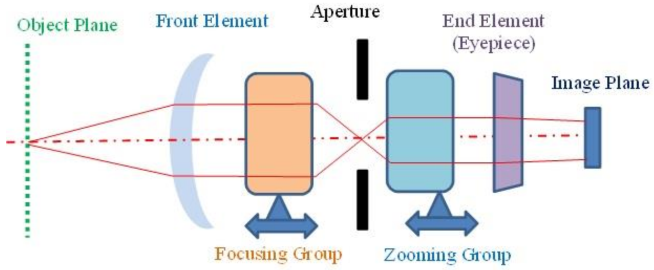


Fig. 1. Zoom Lens generalized structure

Zoom lens elements are grouped to provide basic functions: focusing group, magnification variation group, correction group and image formation group. The magnification variation group and correction group could be moved during zooming. Focusing group could be moved during focusing process that could be performed simultaneously or follow on zooming. Image formation group is usually fixed and belong to the end piece of lens, providing enough space to image plane where imaging sensor is placed (back focus distance).

Back focus distance is the distance along optical axis from the apex of the rearmost glass surface to the focal plane when a lens focused to infinity.

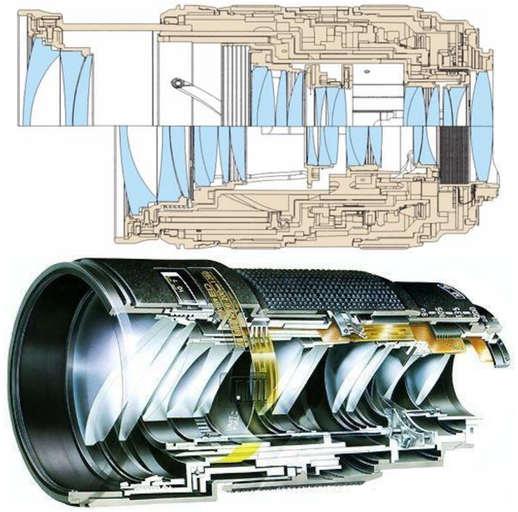


Fig. 2. Zoom lens design example illustration (cross section)

The advances in the optical design and manufacturing process [19], provides high quality for manufactured zoom lenses, but there are still imperfections that remain and that should be corrected during initial lens calibration. Some of them are caused by zoom lens mounting, as illustrated in Fig.3

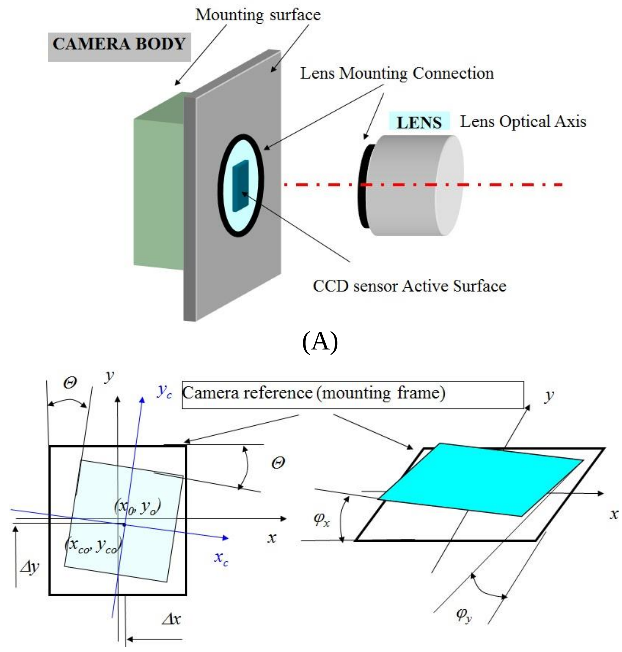


Fig. 3. Zoom lens mounting (A) Back focal distance; (B) Imperfections (Image sensor rotation and tilt)

In the case of surveillance cameras the position image sensor center and optical axis are the most important [20].

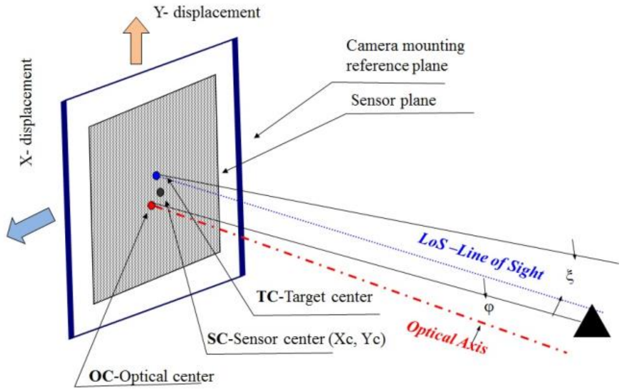


Fig. 4. Zoom lens camera calibration parameters geometry

The fusion of digital image, zoom lens magnification and focus digital control, provides the basis for development of the different calibration procedures that contribute to the effective zoom lens operation in related applications [20-25].

In this article we are dealing with zoom lens calibration to provide reasonably sharp image (according to allowable tolerances of focus [26]) during zoom focal length change but avoiding manual focus fine adjustment. This is of importance in the long range surveillance system application when fast reaction increase system efficacy.

Motorized zoom lenses have great potential in the applications long range surveillance systems providing object of interest - target and visual perception (detection, recognition, identification) and tracking. In such applications, the aperture, zoom, and focus of the lens can be controlled to adapt to different lighting conditions or to obtain the desired field of view, depth of field, spatial resolution, or focused distance. Calibration of a motorized zoom lens is extremely useful but not an easy job. The goal of motorized zoom lens calibration is to determine the relationship between the lens position settings (control parameters for the driving motors) to provide proper imaging performances. In our case we are establishing lens position settings for selected pre-set positions to provide sharp image immediately after zoom position change. In that case zoom lens is treated as mono-focal lens for selected FOV.

III. Hyper-focal Distance and Depth of Field

To provide calibration and set-up procedure we selected to explore the idea of hyper-focal distance that was widely used in first half of twentieth century for focus setting in photography and cinema industry [10, 27]. It is still explored in photography for achieving of the special effects in artistic photography.

Hyper-focal distance can be defined as distance in front of camera upon which all objects in front of that distance and up to infinity appears sharp in image. In other words, if you want everything to be acceptably sharp from infinity to as close as possible to the camera you should focus on the hyper-focal distance.

As a first step we need to discuss some important parameters:

Stop diaphragm - aperture: The opening which adjusts the diameter of the group of light rays passing through the lens. The zoom lens maximal aperture is limited by lens entrance element diameter. To provide possibility to regulate amount of the light the controllable stop diaphragm – iris is built in lens system to limit light bundle passing through lens but not disturbing image. In modern digital cameras aperture adjustment is commonly controlled automatically using command from image sensor to avoid sensor saturation. The control could be complex because saturation is controlled by exposure time, too.

Focal length: When parallel light rays enter the lens parallel to the optical axis, the distance along the optical axis from the lens' second principal point (rear nodal point) to the focal point is called the focal length. In simpler terms, the focal length of a lens is the distance along the optical axis from the lens' second principal point to the focal plane when the lens is focused at infinity.

F-number – F#: For zoom lens having aperture with diameter d and focal length f related lens F-number is defined as:

$$F_{\#} = \frac{f}{d} \tag{1}$$

Circle of confusion – CoC: Since all lenses contain a certain amount of spherical aberration and astigmatism, they cannot perfectly collect rays from a subject point to form a true image point. In addition, diffraction widening also contributes to the focused spot size. In the case of infrared lenses diffraction limited lenses are more often case. Since the image becomes less sharp as the size of these dots increases, the dots are called "circles of confusion". The advancement of the optical design and manufacturing technology practically eliminate the aberrations influence on CoC through proper compensation and optimization in design process, but diffraction influence could not be eliminated.

Diffraction limited resolution is defined by so called Rayleigh's criterion [28, 29] defined by diameter of the first Airy disk on point spread function. In that case one can to define that minimal value of CoC is equal to:

$$c = CoC = 2{,}44 \cdot \lambda \cdot F_{\#} \tag{2}$$

Sensor size: Imaging focal plane array contains the matrix of detectors – pixels having limited size, p. The final factor determining hyper-focal distance is the size of your digital sensor. A larger digital sensor will result in a closer hyper-focal distance.

Depth of focus is a space around focal plane in which acceptable sharp image could be achieved. Depth of focus in the image-space is related to how the quality of focus changes on the sensor side of the lens as the sensor is moved, while the object remains in the same position. Depth of focus, δ, could be calculated as:

$$\delta = 2 \cdot F_{\#} \cdot s \tag{3}$$

Where s *is* resolution limit ($s=p$ – detector limited, $s=c$ – diffraction limited)

Depth of focus characterizes how much tip and tilt is tolerated between the lens image plane and the sensor plane itself. As $F_\#$ decreases, the depth of focus does as well, which increases the impact that tilt has on achieving best focus across the sensor. Without active alignment, there will always be some degree of variation in the orthogonality between the sensor and the lens that is used. It is generally assumed that problems involving depth of focus only occur with large sensors.

Depth of field – DoF: The space in front of and behind an imaged object having sharp focus, image also appears sharp. In other words, the depth of sharpness to the front and rear of the subject where image blur in the focal plane falls within the limits of the permissible circle of confusion. Depth of field (as illustrated in Fig.5) varies according to the lens' focal length, aperture value and object distance (D_{pof}), so if these values are known, a rough estimate of the depth of field can be calculated using the following formulas:

$$D_{near} = \frac{f^2 \cdot D_{pof}}{f^2 + F_\# \cdot c \cdot (D_{pof} - f)} \tag{4}$$

$$D_{far} = \frac{f^2 \cdot D_{pof}}{f^2 - F_\# \cdot c \cdot (D_{pof} - f)} \tag{5}$$

$$DoF = D_{far} - D_{near} \tag{6}$$

$$D_{HF} = \frac{f^2}{F_\# \cdot c} \tag{7}$$

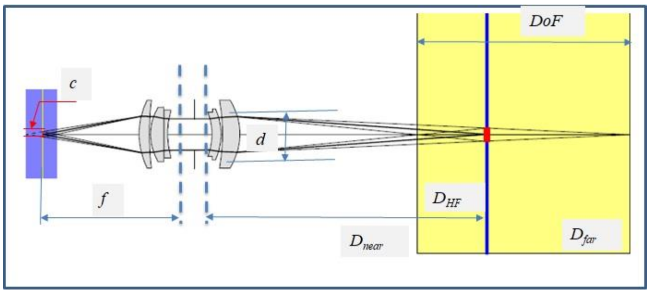


Fig. 5. Depth of focus and depth of field definition

Hyper-focal distance is a distance of object whose image is sharp same as all other far objects up to infinity and other near objects up to half of hyper-focal distance. It depends of the same three factors that determine depth of field.

Hyper-focal distance allows precise setting of the focus so that everything between half the hyper-focal distance and infinity is acceptably sharp. That means if one set proper focus of object placed on hyper-focal distance then all objects placed at distances from half of hyper-focal distance to infinity will have acceptable sharpness.

However, sharpness does not depend on focus setting alone. Camera motion, subject motion by wind, quality of the lens, weather, and other factors can greatly impact the sharpness of the image.

IV. Long Range Zoom Camera Focus Settings

The calculations of the hyper focal distance against focal length setting using equitation (7) were applied for selected cameras having characteristics listed in TABLE I.

TABLE I
Condensed review of the imager basic parameters

System Type	Focal Length [mm]		F#		Pixel Size [µm]	Center λ [µm]
	min	max	min	max		
VOXI-LWIR	25	225	1,5	1,5	17	10
VOXI-SWIR	25	300	2,8	>30*	15	1,5
VOXI-VIS	23	506	3,1	>30*	6,25	0,55
C225	23	506	3,1	>30*	8,3	0,55
C1200	12	1680	4	18,2	5	0,55
C2500	33,4	2000	3,5	16	5	0,55

Note: * - *with iris control*

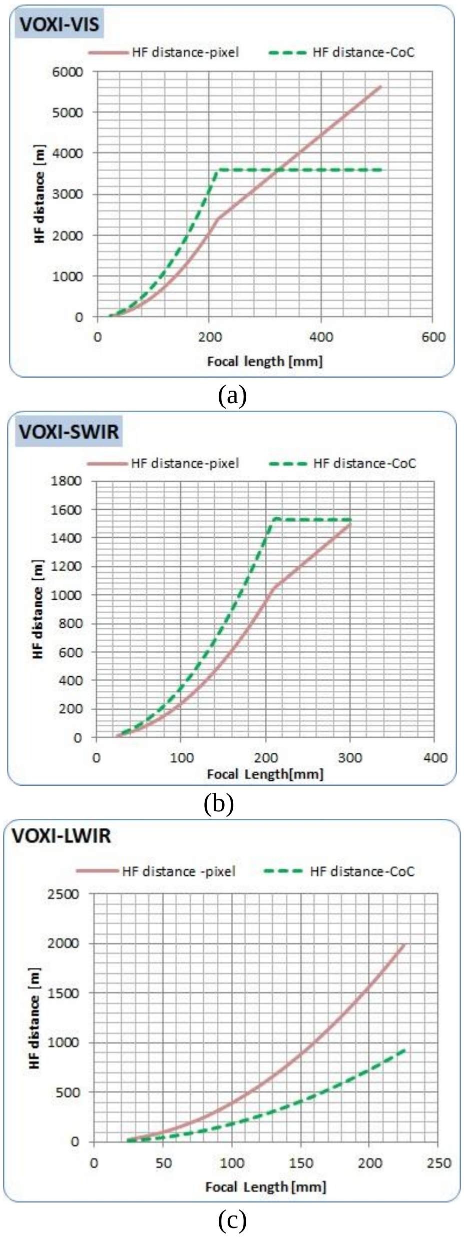


Fig. 6. Calculated hyper-focal distance value versus lens focal length for different cameras (VIS –visible, SWIR- short wavelength infrared, LWIR – long wavelength) in VOXI system

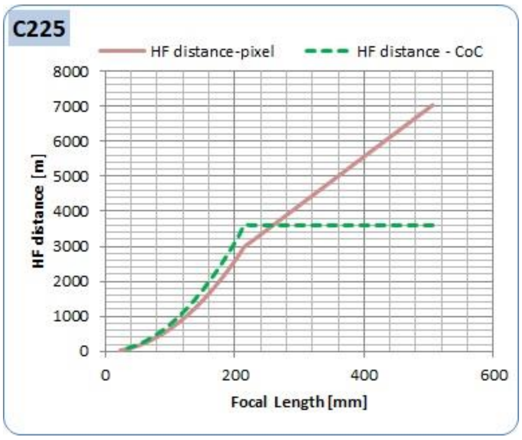


Fig. 7. Calculated hyper-focal distance value versus lens focal length for VIS camera in C225 system

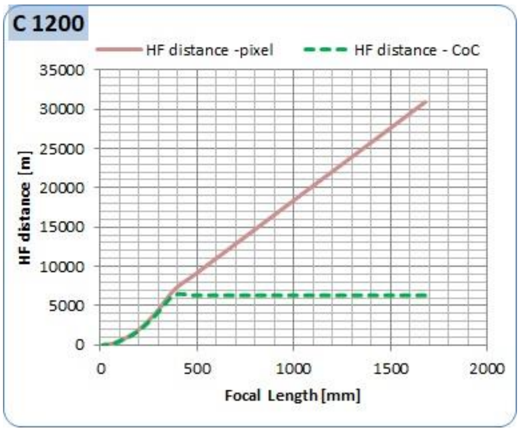


Fig. 8. Calculated hyper-focal distance value versus lens focal length for VIS camera in C1200 system

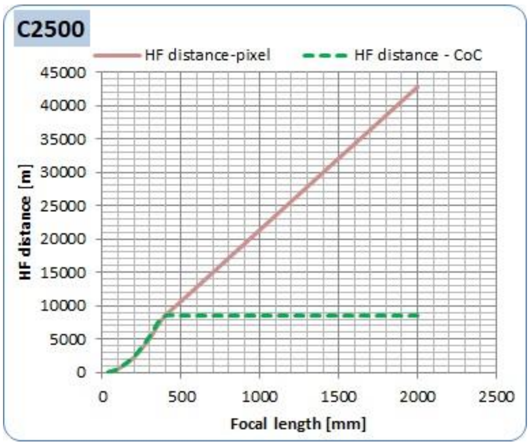


Fig. 9. Calculated hyper-focal distance value versus lens focal length for VIS camera in C2500 system

The results of hyper-focal distance calculations for medium range surveillance system and different imagers are shown in Fig.6. These results show that application of proper set-up will be effective in all three channels.

The results of hyper-focal distance calculations for long range systems (VIS imagers with different magnification) are shown in Fig 7 to Fig.9. These results show that hyper-focal distance based focusing setting could be useful for wide FOVs (focal lengths up to 400 mm) but for narrow FOV it is good enough to use just infinity focus settings.

The application of hyper-focal distance based zoom lens focusing parameter settings can contribute to improvements to system performances but have limited capabilities: (a) calibration process could be applied only for limited and predefined pre-set FOV positions; (b) it is effectively applicable only for zoom lens wide FOV pre-set positions; (c) practical implementation is not easy and reliable enough; (d) the influence of temperature on zoom lens focus stability is not involved in the procedure.

The good news is that additional manual focus setting is always allowable to system operator to try to improve image sharpness.

The application of collimator for focus setting following hyper-focal distance data is not practical in the case of long range imagers, so determination of the imager focus setting should be performed using real objects imaging in the field.

The setting procedure could be established as follows: (I) select the object in the field on the distance that is approximately as calculated hyper-focal distance for selected pre-set FOV (focal length); (II) starting from infinity focusing settings, change the focus setting to get selected object properly focused; (III) continue changing focus setting keeping object focused same as all other objects far from selected object; (IV) at the moment when very far objects (infinity) start to loose sharpness but selected object is still sharp record focus settings. This value could be set as selected focus setting for the FOV pre-set value according to hyper-focal distance.

V. Conclusion

Video channel functional parameter optimization is not simple task, but possibilities depends on image formation process knowledge including lens properties, imaging sensors properties and incorporated camera controls including image processing algorithms.

Application of the zoom lenses in the long range imaging systems is convenient for search and track task over all ranges but diffraction effects at high zoom degrades imaging sensor resolution. Image blurring is inherent to high magnification imaging systems due to lower contrast and high level of diffraction and atmospheric influences. Initial zoom lens set up can help but problem could be solved only by additional image processing algorithm application.

Motorized and controllable zoom lenses together with digital image allow wide range of computer application for application of digital controlled calibration procedures contributing to image quality/sharpness improvements. Camera focusing parameter set-up using hyper-focal distance is calibration process.

The application of the hyper-focal distance concept for camera initial set up provides faster and more comfort operation, but could not contribute alone to image quality/sharpness improvement.

Digital image quality and sharpness depends on lot of factors and could be improved by application of the de-blurring algorithms using the knowledge about physical process causing image blur.

ACKNOWLEDGMENT

We kindly appreciate our colleague Saša Milićević, passionate photographer, who proposed the idea to exploite hyper-focal distance for system focus set-up.

REFERENCES

[1] Shirazi Muhammad Ayaz, Min Young Kim, Jaechan Park, Survey on zoom-lens calibration methods and techniques", *Machine Vision and Applications*, July 2017

[2] Y.Seo and K.S.Hong, "Theory and practice on the self-calibration of a rotating and zooming camera from two views", *IEE Proc.-Vis. Image Signal Process., Vol. 148, No. 3,* June 2001

[3] Andrea Pennisi, Domenico Bloisi, Claudio Gaz, Luca Iocchi, Daniele Nardi, "Novel Patterns and Methods for Zooming Camera Calibration", *Journal of WSCG*, Volume 21, Number 1, 2013, pp, 59-67

[4] Fabio Remondino, Clive Fraser, "Digital Camera Calibration Methods: Consideration and Comparisons", *IAPRS Volume XXXVI, Part 5,* Dresden 25-27 September 2006

[5] S.S. Al-Ajlouni , C.S. Fraser, "Zoom- dependent Calibration for Consumer Grade Cameras", *IAPRS Volume XXXVI, Part 5*, Dresden 25-27 September 2006

[6] Reg G. Willson "Modeling and Calibration of Automated Zoom Lenses", *Proceedings SPIE Volume 2350*, Videometrics III; Photonics for Industrial Applications, Boston, MA, United State, 6 October 1994

[7] Suchi Upadhyay, S.K.Singh, Manoj Gupta, Ashok K. Nagawat, "Linear and Non-linear Camera Calibration Techniques", *Journal of Global Research in Computer Science*, Volume 2, No. 5, April 2011

[8] Bo Wu, Han Hu, Qing Zhu, and Yeting Zhang, "A Flexible Method for Zoom Lens Calibration and Modeling Using a Planar Checkerboard", *Photogrammetric Engineering & Remote Sensing*, Vol. 79, No. 6, June 2013, pp. 555–57

[9] Branko Tomić, Dragana Perić, Miloš Radisavljević, Saša Vujić, „Reticle Integration and Camera Calibration in Multi Sensor Surveillance Systems", *8th International Conference on Defensive Technologies, OTEH 2018,* Belgrade, Serbia, 11-12 October 2018

[10] Rowlands, D. Andrew, *Field guide to photographic science,* SPIE Press, Bellingham, USA, 2020

[11] Allen Mann : "*Infrared Optics and Zoom Lenses"*, Vol. TT42, SPIE Press, Bellingham, USA, 2009

[12] A.D Clark; "*Zoom Lenses*", Monographs in Applied Optics No7, Adam Holger, London, 1973

[13] K. Yamaji: "*Design of Zoom Lenses*" ch.4 in "Progress in Optics" vol.6 (edited by E. Wolf), NORTH-HOLLAND PUBLISHIXG COMPANY, Amsterdam, 1967

[14] Richard N.Youngworth, Ellis I. Betensky: "Fundamental Considerations for Zoom lens Design (Tutorial)", *Proc. of SPIE,* Vol. **8488** p. 848806-1, 2012

[15] Allen Mann: "Infrared zoom lens system for target detection", Optical Engineering, Vol. 21 No. 4, pp.786-793, 1982

[16] Thomas H. Jamieson: „Thin-lens theory of zoom systems", OPTICA ACTA, VOL. 17, NO. 8, 565—584, 1970

[17] Antonín Mikš, Jiří Novák, and Pavel Novak: "Method of zoom lens design", *Applied Optics*, Vol. 47, No. 32, pp. 6088-6098, 2008

[18] D. Dilworth: "A zoom lens from scratch: the case for number crunching", *Proc. SPIE*. 9947, 994702, 2016

[19] Robert E. Fischer, Biljana Tadic-Galeb, Paul R. Yoder: "*Optical System Design*", The McGraw-Hill Companies, Inc., New York, 2008

[20] Reg G. Willson, Steven A. Shafer, "What is the center of the image?", *J. Opt. Soc. Am. A* ,/ Vol. 11, No. 11/, November 1994

[21] Zheng Wang, Jon Mills, Wen Xiao, Rongyong Huang, Shunyi Zheng and Zhenhong Li, "A Flexible, Generic Photogrammetric Approach to Zoom Lens Calibration", *Remote Sens.* 2017, 9, 244

[22] Yong-Sheng Chen Sheng-Wen Shih, Yi-Ping Hung, Chiou-Shann Fuh, "Camera Calibration with a Motorized Zoom Lens", *Proceedings of the 15th International Conference on Pattern Recognition*, Barcelona, Spain, Sep. 2000, Volume 4, pages 495–498

[23] Anthony G. Wlley and Kam W. Wong, "Geometrical Calibration of Zoom Lenses for Computer Vision Metrology", *Photogrammetric Engineering & Remote Sensing*, Vol. 61, No. 1, January 1995, pp. 69-74

[24] Yong-Sheng Chen, Sheng-Wen Shih, Yi-Ping Hung, Chiou-Shann Fuh: " Simple and efficient method of calibration a motorized zoom lens", *Image and Vision Computing*, 19, 1099-1110, 2001

[25] Ziyan Wu, and Richard J. Radke: „Keeping a Pan-Tilt-Zoom Camera Calibrated", *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, pp 1-14, 2012

[26] S.G. Shiue, „Tolerances for Focus Stability in Zoom Lenses", *Optica Acta: International Journal of Optics,* 32:8, 923-936, 1985

[27] Sidnay F. Ray: "*Applied Photographic Optics: Lenses and optical systems for photography, film, video, electronic and digital imaging*", third edition, Focal Press, Oxford, 2002

[28] A. J. den Dekker and A. van den Bos: „Resolution: a survey", *J. Opt. Soc. Am. A* , Vol. 14, No. 3, 547-557 , 1997

[29] Sisi Zhou and Liang Jiang: "A modern description of Rayleigh's criterion", *Phys. Rev. A* , 99, 013808, 2019

Temperature influence on the performance of P3HT:ICBA polymer solar cells

Ali R. Khalf, Jovana P. Gojanović, Nataša A. Ćirović, Petar S. Matavulj, *Member, IEEE*, Grant Ledet, Mark Hidalgo and Sandra Živanović, *Member, IEEE*

***Abstract*—Temperature (*T*) dependent performance of polymer solar cells (PSCs) with a poly (3-hexylthiophene): indene-C_{60} bisadduct (P3HT:ICBA) active layer were investigated. The current-voltage *(I-V)* characteristics of devices with two different active layer thicknesses (ALTs) were measured within a temperature range of 20 °C-65 °C. The recorded *I-V* curves showed the S-shape deviation. The *I-V* curves were also simulated by a standard drift-diffusion model that includes the influence of the surface recombination on both electrode contacts. The Arrhenius-type temperature-dependent hole mobility was introduced to reproduce the experimentally observed temperature-dependent PSC behavior. The measured power conversion efficiency (PCE) and short-circuit current (I_{sc}) changed non-monotonically with *T* whereby detailed temperature trends differed for solar cells of different ALTs. The noticed effects were not present in theoretically predicted PCE and I_{sc}. To match the simulated and experimental *I-V* characteristics the PCS internal quantum efficiency (IQE) was varied with *T*. We suggest that the obtained nonmonotonic IQE(*T*) dependence originates from changes in morphology caused by the influence of temperature and strongly correlates to the P3HT:ICBA thin-film thickness.**

***Index Terms*—P3HT; solar cells; simulation; temperature.**

I. Introduction

POLYMER-based bulk-heterojunction solar cells (PSCs) are an emerging renewable energy technology that enables easy, low-cost and low-environmental impact production and yields lightweight, flexible devices with the possibility of visible transparency, and large surface area. Power conversion efficiency (PCE) of PSCs has drastically improved during the last decade, surpassing 18% for single-junction cells [1], and 18.6% for tandem cells [2]. Enhancement of PCE has been accomplished through several different developmental directions. A crucial issue regarding polymer-based solar cells is how to manage the energy levels of the donor/acceptor (D/A) blends to maximize short-circuit current (I_{sc}) and open-circuit voltage (V_{oc}) at the same time without sacrificing the efficient charge separation [3, 4]. New donor [3, 4] and acceptor [5] materials were synthesized to accomplish this goal. Optimization of process parameters [6], annealing [7], aggregation, and morphology control [8] were used to improve the charge transport in PSCs. Much better extraction of charge carriers was accomplished by introducing the hole and electron buffer layers [9]. Optical manipulation of light has become an increasingly popular strategy to capture solar radiation more effectively into an ultrathin photoactive layer of PSC thus to enhance the light-harvesting efficiency [10, 11].

In the meantime, a lot of theoretical research has been done resulting in the first PSC drift-diffusion model (DDM) developed by Koster et al. in 2005 [12]. A long time ago, it was established that interferential effects play a significant role in organic thin-film photovoltaic devices [13]. Therefore, an optical model based on transfer matrix formalism was soon coupled to drift-diffusion calculations, completing the image of important physical processes in PSC [13]. This model including different modifications and updates is successfully used to simulate the performance of PSC with various structures and D/A material combinations. It is also a powerful tool for the investigation of physical phenomena that undergo PSC operation [14, 15] as well as for device optimization [16]. Besides the drift-diffusion approach, some equivalent circuit models have also been proposed [17]. These models introduce the other point of view, and they account for additional electrical PSC features not included in the DDM such as parasitic resistivities and other parasitic effects. Another field that the DDM does not cover in a sufficiently detailed way is the impact of morphology and nanoscale physical processes on the efficiency of PSC. Through Monte Carlo and multiscale simulations [18], one can approach the nanostructure of the active layer and follow the excitonic and charge carrier pathways. This can lead to some crucial conclusions and hints for the fabrication of highly efficient PSC.

Summarizing the state of the art in the field of PSCs, it becomes clear that there is a lot of room for additional research by physicists, chemists, and technologists to improve PSC efficiency towards their commercialization.

It is well known that temperature and light intensity dependence of optoelectronic device performance gives a good insight into the physical processes underlying its operation. Such measurements were carried out on the PSCs to study the photogeneration and transport of charge carriers [19, 14, 15, 20 21] as well as the mechanisms of their recombination [22, 23, 24, 20]. According to our knowledge only a few papers in the literature have been dedicated to the investigation of PSC *I-V* curve temperature-dependence and consequently I_{sc}, V_{oc}, fill factor (FF), and PCE temperature dependences [25, 26, 27, 28, 29, 30, 20]. Among these papers just one presents the DDM model that includes the PCS temperature dependent behavior, unfortunately, without matching the model results to any experimental data [20]. Knowing the influence of the temperature on PCS performance is very important to predict the operation of the device in standard working conditions as well as for further progress in PCE optimization. To prevent unnecessary

A. R. Khalf, J. P. Gojanović, N. A. Ćirović, and P. S. Matavulj are with the School of Electrical Engineering, University of Belgrade, Belgrade 11120, Serbia (e-mail: alirkhalf@yahoo.com; jovana@etf.bg.ac.rs; natasa@etf.bg.ac.rs; matavulj@etf.bg.ac.rs).

G. Ledet, M. Hidalgo and S. Živanović are with Institute for Micromanufacturing, Louisiana Tech University, Ruston, LA 71272 USA (e-mail: grantledet@gmail.com, mkhidalgo@charter.net, sz@latech.edu).

expenses, a useful model is needed to simulate temperature-dependence of PSC.

In this paper ITO(indium tin oxide)/PEDOT:PSS (poly(3,4-hylenedioxythiophene):poly(styrenesulfonate))/P3HT:ICBA (poly(3-hexylthiophene):indene-C_{60} bisadduct)/Al solar cells with two different thicknesses of P3HT:ICBA active layer were fabricated and tested in the 20 °C to 65 °C temperature range. The S-shaped *I-V* characteristics were recorded under solar simulator light from which I_{sc} and PCE dependences on T were determined for each device. The device performance was simulated by the standard drift-diffusion model (DDM) that accounts for surface recombination on electrodes. The Arrhenius-type temperature-dependent hole mobility was applied [15]. It was found that the experimental PCE and I_{sc} vary non-monotonicaly with temperature (T). Also, temperature caused variation of PCE and I_{sc} for two devices with different active layer thicknesses (ALTs) was different. The DDM calculated solar cell parameters did not show such behavior. To match the simulated and measured *I-V* curves it was necessary to introduce the temperature-dependent internal quantum efficiency (IQE) of PSCs. The surface recombination was also taken to be temperature-dependent. The nonmonotonic change of IQE with T obtained in this way was attributed to the change of P3HT:ICBA film morphology which is, on the other hand, correlated to the active layer film thickness [31].

II. EXPERIMENT

PSC devices with glass/ ITO/ PEDOT: PSS/ P3HT:ICBA/ Al device structure were fabricated and tested at Institute for Micromanufacturing, Louisiana Tech University.

P3HT and ICBA from Sigma Aldrich with 1:0.78 wt. ratio were mixed with chlorobenzene separately and kept on a hot plate with a magnetic stirrer at 50 °C overnight. PEDOT:PSS HTL Solar Heraeus Clevios water solution from Ossila was spin-coated at 3500 RPM to deposit about 50 nm-thick film and was then annealed at 115°C for 10 min. The P3HT:ICBA solution was then statically dispensed with a micropipette onto the ITO coated boro-aluminosilicate glass substrate (Delta Technologies) and spin-coated at 900 RPM and 1450 RPM to deposit approximately 90 nm and 67 nm-thick films, respectively. The P3HT:ICBA thin films were baked at 70 °C for 5 minutes to remove any residual solvent. Afterwards, a 110 nm aluminum layer was deposited as a cathode in the e-beam evaporator, and then the devices were annealed at 150 °C for 15 min on a conventional hot plate. The active area of each device was about 0.9 cm^2.

Devices were illuminated with AM1.5 spectra of 50-60 mW/cm^2 optical power density from Spectra Physics 66900 solar simulator. The incident optical power density was measured with a Newport Oriel 91150V reference cell and meter. To control the temperature of the solar cells for testing, a thermoelectric Peltier module with a DC voltage supply, which uses voltage to change the temperature on the plates, was used. A non-contact infrared thermometer was also used to monitor the temperature. The *I-V* curves were measured for the temperature range from 20 °C to 65 °C using Keithley 2400 source meter. Fig. 1 (a_1) and (b_1) show experimentally obtained *I-V* curves under solar simulator illumination. From the *I-V* characteristics, the corresponding temperature-dependent I_{sc} and PCE were determined (Fig. 1 (a_2), (a_3), (b_2), and (b_3)).

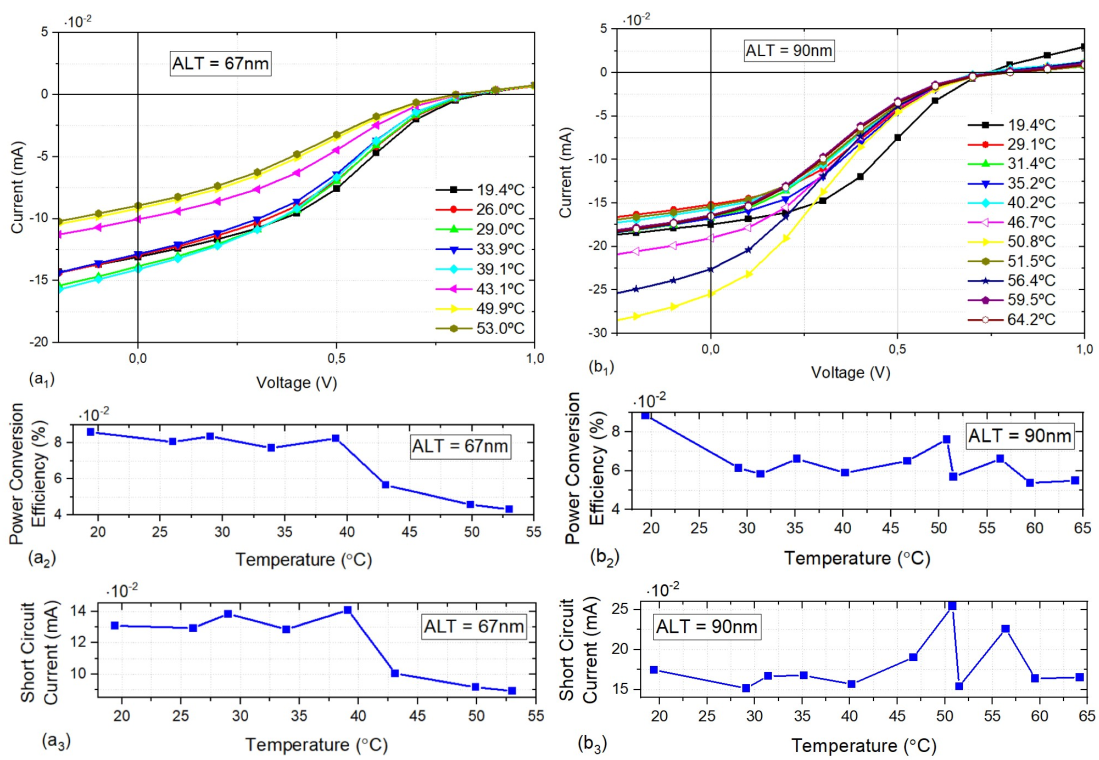


Fig. 1. Measured I-V characteristics of ITO/PEDOT:PSS/P3HT:ICBA/al solar cells with (a1) 67 nm and (b1) 90 nm active layer thicknesses at different Temperatures; (a2) and (b2) PCE and (a3) and (b3) Isc temperature dependencies for the same devices, respectively.

III. Model

To simulate the temperature dependence of ITO/PEDOT:PSS/P3HT:ICBA/Al solar cell characteristics, we used the DDM described in our previous research [32]. The Robin type boundary conditions which account for surface recombination on electrode contacts were applied [33]. Generation rate of charge carriers in the active layer was calculated using the transfer matrix method (TMM) which takes into account interference effects in the device [13]. The optical constants, refraction index, and extinction coefficient used in the TMM were determined from optical measurements that also take the interference effects in the thin organic films into account. As the transport mechanism of holes and electrons in polymer:fullerene blends are strongly thermally activated [14, 15], we supposed that temperature-dependent behaviour of ITO/PEDOT:PSS/P3HT:ICBA/Al solar cell is predominantly governed by temperature-dependent mobilities. Since mobility temperature dependence is much weaker for electrons than for holes [29], we assumed a constant electron mobility and an Arrhenius-type [15] temperature-dependent hole mobility:

$$\mu_p = \mu_{p0} e^{-\frac{\Delta}{k_b T}} \tag{1}$$

where μ_{p0} is a mobility prefactor, Δ is the activation energy, k_b Boltzmann's constant, and T the absolute temperature. The numerical DDM solution includes the Sharfetter-Gummel approach. All the other model methods and assumptions are the same as in [32].

IV. Results and discussion

The *I-V* characteristics for two ITO/PEDOT:PSS/P3HT:ICBA/Al devices with 67 nm and 90 nm thick P3HT:ICBA active layer measured at several different temperatures between 20 °C to 65 °C are shown in Fig. 1 a_1) and b_1). Both devices exhibit pronounced S-shape deformation. The deformation most likely originates from the aggravated extraction through the Al electrode and a consequent accumulation of the charge carriers. The PCE and I_{sc} as functions of T for two examined PSCs are presented in Fig. 1 a_2, a_3, b_2, and b_3. The non-mnotonic change of PCE and I_{sc} with temperature is apparent for both devices. While the PCE(T) had an overall decreasing character, the maximum of I_{sc} was obtained around 40 °C for the solar cell with 67 nm ALT and around 50 °C for the solar cell with 90 nm ALT. The tested devices had rather small PCE values due to inefficient cathode extraction.

The DDM simulations were conducted by using the parameter values given in Table 1. Because of the S-shape experimental *I-V* characteristics anomaly, it was proposed that surface recombination of electrons was pronounced at the cathode contact. The surface recombination velocities (SRVs) on the anode contact were taken to be infinite, while the electron SRV on the cathode was assumed to be reduced (Table 1). The calculated *I-V* curves for the 67 and 90nm ALT devices for the same T values at which the measurements were done are shown in Fig. 2 (a_1) and (b_1). The corresponding PCE(T) and I_{sc}(T) were determined and depicted in Fig. 2 (a_2), (a_3), (b_2), and (b_3). From Fig. 2 (a_2) and (b_2) it can be noticed that theory predicts a slightly increasing trend for PCE(T). Also, the increasing trend for I_{sc}(T) is obtained from calculations as can be seen from Fig. 2 (a_3) and (b_3). Since the experimental results given in Fig. 1 are qualitatively poorly reproduced by the model, a conclusion was drawn that the DDM which includes the effect of temperature on the PSC performance only through Arrhenius T-dependent hole transport is not adequate. Apparently, there are other processes that are significantly affected by T.

TABLE 1
The parameters for ITO/PEDOT:PSS/P3HT:ICBA/Al solar cells

Symbol	Quantity	Value
E_g	Energy gap	1.4eV
ε_r	Relative permittivity	3.4
N_C, N_V	The effective densities of states	$1\times10^{26}\mathrm{m}^{-3}$
μ_{p0}	Hole mobility	$3\ \mathrm{cm^2/(V{\cdot}s)}$
μ_n	Electron mobility	$4.92\times10^{-4}\ \mathrm{cm^2/(V{\cdot}s)}$
Δ	Activation energy	0.3eV
IQE	Internal quantum efficiency	0.05
τ_n	Electron lifetime	6.2×10^{-5}s
τ_p	Hole lifetime	3×10^{-7}s
S_n^c	Electron SRV at the cathode	9.6×10^{4} cm/s

When a photon is absorbed in the PSC's active layer, the Coulombically bound electron-hole pair constituting an excitonic state is produced. It is known that temperature plays an important role in exciton dissociation [19, 34]. First, the charge transfer (CT) state is made at the donor/acceptor junction and some additional energy is needed to complete the dissociation. Thermal energy can be used for the separation of the CT state into free carriers [19, 34]. Knowing that the photogeneration in PSCs is affected by temperature, we attempted to reproduce the PSCs *I-V* characteristics at different T by letting the IQE be T-dependent and used it as a fitting parameter. The electron SRV on the cathode is also expected to change with T since it is essentially the Shockley-Read-Hall recombination through the surface trap states. For this reason, the SRV for electrons on the cathode was taken to be variable. When the IQE and electron SRV on the cathode were changed with T, a very good agreement between experimental and simulated *I-V* data was accomplished. The comparison of measured and DDM calculated *I-V* curves for solar cells with 67 nm and 90 nm ALTs for three selected temperatures is shown in Fig. 3. The built-in voltage was also slightly changed with temperature, which is denoted on each *I-V* graph. The IQE and SRV values at different T obtained for the two considered PSCs are presented in Table 2. The IQE changed with T in a nonmonotonic fashion, which was not the same for devices with different ALTs. This can be attributed to the P3HT:ICBA film morphology changes caused by the temperature and correlated to the film thickness [31]. On the other hand, the electron SRVs on the cathodes for both devices were decreasing with T, reflecting the fact that the surface (trap-assisted) recombination becomes more pronounced with increasing T.

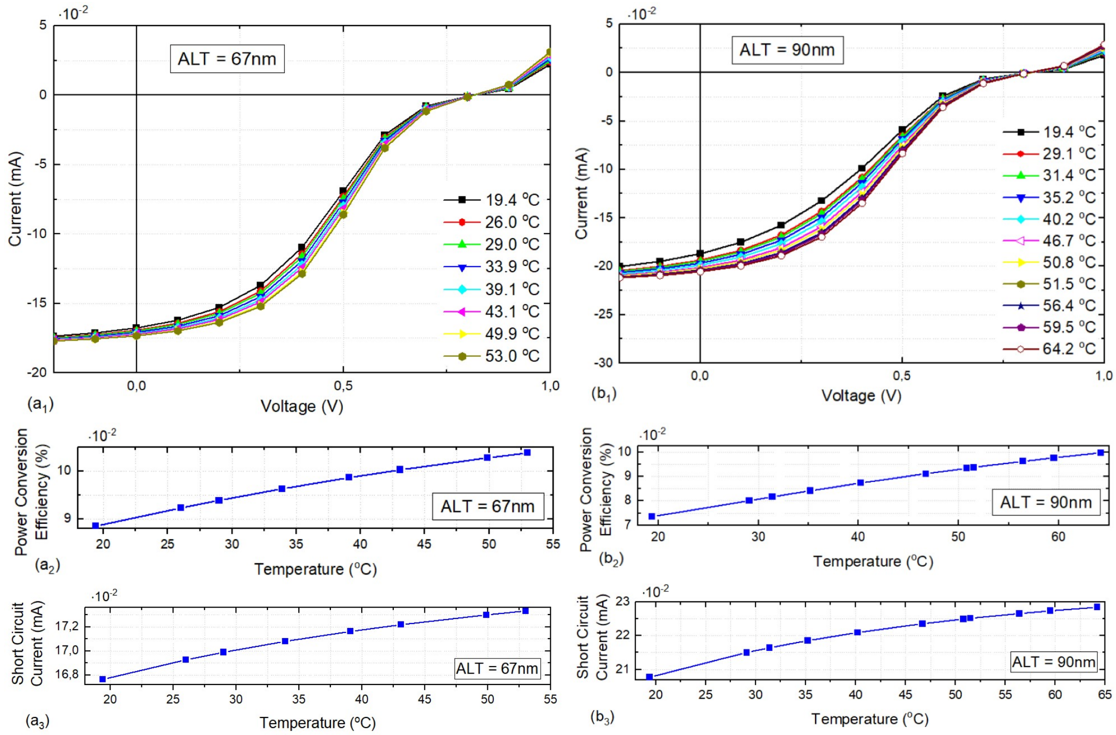


Fig. 2. The I-V characteristics of ITO/PEDOT:PSS/P3HT:ICBA/al solar cells with (a1) 67 nm and (b1) 90 nm active layer thicknesses at different temperatures simulated by standard DDM including Arrhenius-type temperature-dependent hole mobility. The calculated (a2) and (b2) PCE and (a3) and (b3) Isc temperature-dependencies for the same devices, respectively.

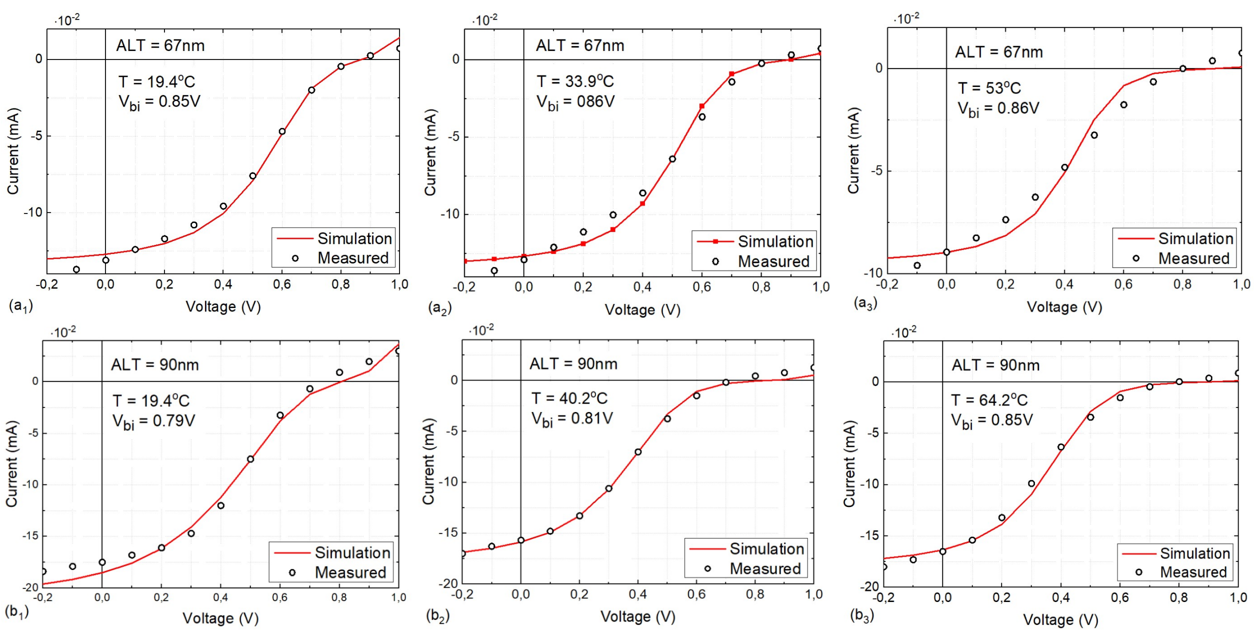


Fig. 3. Comparison of measured and simulated I-V characteristics of ITO/PEDOT:PSS/P3HT:ICBA/al solar cells with 67 nm and 90 nm active layer thicknesses at selected temperatures. Simulations are conducted using DDM with Arrhenius-type temperature-dependent hole mobility together with temperature-dependent IQE and electron SRV at the cathode.

V. CONCLUSION

Temperature influence on the performance of P3HT:ICBA solar cells was investigated. ITO/PEDOT:PSS/P3HT:ICBA/Al solar cells with two different ALTs were fabricated and characterized under solar simulator light. The *I-V* characteristics were measured within the temperature range of 20 °C-65 °C for both devices, and PCE and I_{sc} as functions of T were determined. The recorded *I-V* curves manifested the S-shape deviation. To simulate the *I-V*(*T*) data, the DDM model including surface recombination on both electrodes and Arrhenius-type *T*-dependent hole mobility was used and resulted in a poor agreement between theory and experiment. To better reproduce the experimental *I-V* curves, the *T*-dependent IQE and electron SRV at cathode were assumed and used in calculations as fitting parameters. This way, a very good match of calculated with measured *I-V* curves was achieved in the whole temperature range. The obtained IQE(T) dependence was nonmonotonic and differed for solar cells with different ALTs. This was attributed to the temperature-induced morphology changes, which are strongly correlated with the P3HT:ICBA thin film thickness [31]. The electron SRV at the cathode showed increasing character with T, which is in accord with the fact that it is a trap-assisted recombination mechanism. A further study should be conducted to resolve the correlation between polymer thin film morphology, thickness, and applied temperature.

TABLE 2
THE IQE AND SRV VALUES AT DIFFERENT T

ALT= 90 nm		
T (°C)	IQE	Electron SRV [cm/s]
19.4	0.044	8.90×10^{-3}
29.1	0.037	4.55×10^{-4}
31.4	0.040	4.56×10^{-4}
35.2	0.040	5.74×10^{-5}
40.2	0.037	4.63×10^{-5}
46.7	0.046	4.68×10^{-5}
50.8	0.059	1.57×10^{-5}
51.5	0.036	1.57×10^{-5}
56.4	0.053	1.58×10^{-5}
59.5	0.037	4.77×10^{-6}
64.2	0.037	4.80×10^{-6}
ALT= 67 nm		
T (°C)	IQE	Electron SRV [cm/s]
19.4	0.037	4.47×10^{-3}
26.0	0.037	6.46×10^{4}
29.0	0.040	3.03×10^{-4}
33.9	0.037	1.99×10^{-4}
39.1	0.040	9.24×10^{-5}
43.1	0.029	4.65×10^{-5}
49.9	0.027	2.35×10^{-5}
53.0	0.026	6.75×10^{-6}

ACKNOWLEDGMENT

This work is partially supported by the Serbian Ministry of Education, Science and Technological Development under Grant #171011 awarded to J. Gojanović and P. Matavulj, the National Aeronautics and Space Administration (NASA) grant and Cooperative Agreement Number NNX15AH82H through LaSpace LURA subaward PO-0000107276, and James W. Adams endowed professorship of S. Živanović that is made available through the State of Louisiana Board of Regents Support Funds.

REFERENCES

[1] M. B. Salim, R. Nekovei, and R. Jeyakumar, "Organic tandem solar cells with 18.6% efficiency," *Solar Energy*, vol. 198, pp. 160–166, Mar. 2020, doi: 10.1016/j.solener.2020.01.042.

[2] Q. Liu, Y. Jiang, K. Jin, J. Qin, J. Xu, W. Li, J. Xiong, J. Liu, Z. Xiao, K. Sun, S. Yang, X. Zhang, and L. Ding, "18% Efficiency organic solar cells," *Science Bulletin*, vol. 65, no. 4, pp. 272–275, Feb. 2020, doi: 10.1016/j.scib.2020.01.001.

[3] H. Chen, J. Hou, S. Zhang, Y. Liang, G. Yang, Y. Yang, L. Yu, Y. Wu and G. Li, "Polymer solar cells with enhanced open-circuit voltage and efficiency," *Nature Photon*, vol. 3, no. 11, pp. 649–653, Nov. 2009, doi: 10.1038/nphoton.2009.192.

[4] K. Kawashima, Y. Tamai, H. Ohkita, I. Osaka, and K. Takimiya, "High-efficiency polymer solar cells with small photon energy loss," *Nat Commun*, vol. 6, no. 1, p. 10085, Dec. 2015, doi: 10.1038/ncomms10085.

[5] H. Kang, W. Lee, J. Oh, T. Kim, C. Lee, and B. J. Kim, "From Fullerene–Polymer to All-Polymer Solar Cells: The Importance of Molecular Packing, Orientation, and Morphology Control," *Acc. Chem. Res.*, vol. 49, no. 11, pp. 2424–2434, Nov. 2016, doi: 10.1021/acs.accounts.6b00347.

[6] S. Jin, B. V. K. Naidu, H. Jeon, S. Park, J. Park, S. C. Kim, J. W. Lee, Y. Gal, "Optimization of process parameters for high-efficiency polymer photovoltaic devices based on P3HT:PCBM system," *Solar Energy Materials and Solar Cells*, vol. 91, no. 13, pp. 1187–1193, Aug. 2007, doi: 10.1016/j.solmat.2007.04.001.

[7] J. G. Sanchez, A. A. A. Torimtubun, V. S. Balderrama, M. Estrada, J. Pallares, and L. F. Marsal, "Effects of Annealing Temperature on the Performance of Organic Solar Cells Based on Polymer: Non-Fullerene Using V_2O_5 as HTL," *IEEE J. Electron Devices Soc.*, vol. 8, pp. 421–428, 2020, doi: 10.1109/JEDS.2020.2964634.

[8] Y. Liu, J. Zhao, Z. Li, C. Mu, W. Ma, H. Hu, K. Jiang, H. Lin, H. Ade, and H. Yan, "Aggregation and morphology control enables multiple cases of high-efficiency polymer solar cells," *Nat Commun*, vol. 5, no. 1, p. 5293, Dec. 2014, doi: 10.1038/ncomms6293.

[9] F. Cheng, G. Fang, X. Fan, H. Huang, Q. Zheng, P. Qin, H. Lei, and Y. Li, "Enhancing the performance of P3HT:ICBA based polymer solar cells using LiF as electron collecting buffer layer and UV–ozone treated MoO3 as hole collecting buffer layer," *Solar Energy Materials and Solar Cells*, vol. 110, pp. 63–68, Mar. 2013, doi: 10.1016/j.solmat.2012.12.006.

[10] S. B. Jo, Y. Liu, F. Lin, and A. K. -Y. Jen, "Tackling Energy Loss for High-Efficiency Organic Solar Cells with Integrated Multiple Strategies," *Adv. Mater.*, vol. 30, no. 16, p. 1706816, Apr. 2018, doi: 10.1002/adma.201706816.

[11] W. Cai, X. Gong, and Y. Cao, "Polymer solar cells: Recent development and possible routes for improvement in the performance," *Solar Energy Materials and Solar Cells*, vol. 94, no. 2, pp. 114–127, Feb. 2010, doi: 10.1016/j.solmat.2009.10.005.

[12] L. J. A. Koster, E. C. P. Smits, V. D. Mihailetchi, and P. W. M. Blom, "Device model for the operation of polymer/fullerene bulk heterojunction solar cells," *Phys. Rev. B*, vol. 72, no. 8, p. 085205, Aug. 2005, doi: 10.1103/PhysRevB.72.085205.

[13] L. A. A. Pettersson, L. S. Roman, and O. Inganäs, "Modeling photocurrent action spectra of photovoltaic devices based on organic thin films," *Journal of Applied Physics*, vol. 86, no. 1, pp. 487–496, Jul. 1999, doi: 10.1063/1.370757.

[14] V. D. Mihailetchi, H. X. Xie, B. de Boer, L. J. A. Koster, and P. W. M. Blom, "Charge Transport and Photocurrent Generation in Poly(3-hexylthiophene): Methanofullerene Bulk-Heterojunction Solar Cells," *Adv. Funct. Mater.*, vol. 16, no. 5, pp. 699–708, Mar. 2006, doi: 10.1002/adfm.200500420.

[15] F. Laquai, D. Andrienko, R. Mauer, and P. W. M. Blom, "Charge Carrier Transport and Photogeneration in P3HT:PCBM Photovoltaic Blends," *Macromol. Rapid Commun.*, vol. 36, no. 11, pp. 1001–1025, Jun. 2015, doi: 10.1002/marc.201500047.

[16] Y. Min Nam, J. Huh, and W. H. Jo, "A computational study on optimal design for organic tandem solar cells," *Solar Energy Materials and Solar Cells*, vol. 95, no. 4, pp. 1095–1101, Apr. 2011, doi: 10.1016/j.solmat.2010.12.018.

[17] A. Cheknane, H. S. Hilal, F. Djeffal, B. Benyoucef, and J.-P. Charles, "An equivalent circuit approach to organic solar cell modelling," *Microelectronics Journal*, vol. 39, no. 10, pp. 1173–1180, Oct. 2008, doi: 10.1016/j.mejo.2008.01.053.

[18] G. Li, L. Liu, F. Wei, S. Xia, and X. Qian, "Recent Progress in Modeling, Simulation, and Optimization of Polymer Solar Cells," IEEE J. Photovoltaics, vol. 2, no. 3, pp. 320–340, Jul. 2012, doi: 10.1109/JPHOTOV.2012.2193385.

[19] C. Ma, C. C. S. Chan, X. Zou, H. Yu, J. Zhang, H. Yan, K. S. Wong, and P. C. Y. Chow, "Unraveling the Temperature Dependence of Exciton Dissociation and Free Charge Generation in Nonfullerene Organic Solar Cells," *Sol. RRL*, vol. 5, no. 4, p. 2000789, Apr. 2021, doi: 10.1002/solr.202000789.

[20] A. Mahmoudloo and S. Ahmadi-Kandjani, "Influence of the temperature on the charge transport and recombination profile in organic bulk heterojunction solar cells: a drift-diffusion study," *Appl. Phys. A*, vol. 119, no. 4, pp. 1523–1529, Jun. 2015, doi: 10.1007/s00339-015-9130-3.

[21] K. Tvingstedt, J. Benduhn, and K. Vandewal, “Temperature dependence of the spectral line-width of charge-transfer state emission in organic solar cells; static *vs.* dynamic disorder,” *Mater. Horiz.*, vol. 7, no. 7, pp. 1888–1900, 2020, doi: 10.1039/D0MH00385A.

[22] S. A. Hawks, G. Li, Y. Yang, and R. A. Street, “Band tail recombination in polymer:fullerene organic solar cells,” *Journal of Applied Physics*, vol. 116, no. 7, p. 074503, Aug. 2014, doi: 10.1063/1.4892869.

[23] S. R. Cowan, N. Banerji, W. L. Leong, and A. J. Heeger, “Charge Formation, Recombination, and Sweep-Out Dynamics in Organic Solar Cells,” *Adv. Funct. Mater.*, vol. 22, no. 6, pp. 1116–1128, Mar. 2012, doi: 10.1002/adfm.201101632.

[24] S. R. Cowan, A. Roy, and A. J. Heeger, “Recombination in polymer-fullerene bulk heterojunction solar cells,” *Phys. Rev. B*, vol. 82, no. 24, p. 245207, Dec. 2010, doi: 10.1103/PhysRevB.82.245207.

[25] A. Petrovic, J. Gojanovic, P. Matavulj, M. Islam, and S. Zivanovic, “Temperature dependence of P3HT:ICBA polymer solar cells,” in *2017 International Conference on Numerical Simulation of Optoelectronic Devices (NUSOD)*, Copenhagen, Denmark, Jul. 2017, pp. 133–134. doi: 10.1109/NUSOD.2017.8010027.

[26] G. Garcia-Belmonte, “Temperature dependence of open-circuit voltage in organic solar cells from generation–recombination kinetic balance,” *Solar Energy Materials and Solar Cells*, vol. 94, no. 12, pp. 2166–2169, Dec. 2010, doi: 10.1016/j.solmat.2010.07.006.

[27] I. Riedel, J. Parisi, V. Dyakonov, L. Lutsen, D. Vanderzande, and J. C. Hummelen, “Effect of Temperature and Illumination on the Electrical Characteristics of Polymer–Fullerene Bulk-Heterojunction Solar Cells,” *Adv. Funct. Mater.*, vol. 14, no. 1, pp. 38–44, Jan. 2004, doi: 10.1002/adfm.200304399.

[28] K. Tvingstedt and C. Deibel, “Temperature Dependence of Ideality Factors in Organic Solar Cells and the Relation to Radiative Efficiency,” *Adv. Energy Mater.*, vol. 6, no. 9, p. 1502230, May 2016, doi: 10.1002/aenm.201502230.

[29] B. Yang, J. Cox, Y. Yuan, F. Guo, and J. Huang, “Increased efficiency of low band gap polymer solar cells at elevated temperature and its origins,” *Appl. Phys. Lett.*, vol. 99, no. 13, p. 133302, Sep. 2011, doi: 10.1063/1.3643450.

[30] W. Bagienski and M. C. Gupta, “Temperature dependence of polymer/fullerene organic solar cells,” *Solar Energy Materials and Solar Cells*, vol. 95, no. 3, pp. 933–941, Mar. 2011, doi: 10.1016/j.solmat.2010.11.026.

[31] J. Vlahović, M. Stanojević, J. Gojanović, J. Melancon, A. Sharma, and S. Živanović, “Thickness dependent photocurrent spectra and current-voltage characteristics of P3HT:PCBM photovoltaic devices,” *Opt. Express*, vol. 29, no. 6, p. 8710, Mar. 2021, doi: 10.1364/OE.418082.

[32] Ž. Jelić, J. Petrović, P. Matavulj, J. Melancon, A. Sharma, C. Zellhofer, and S. Živanović, “Modeling of the polymer solar cell with a P3HT:PCBM active layer,” Phys. Scr., vol. T162, p. 014035, Sep. 2014, doi: 10.1088/0031-8949/2014/T162/014035.

[33] A. R. Khalf, J. P. Gojanovic, N. A. Cirovic, S. Zivanovic, and P. S. Matavulj, “The Impact of Surface Processes on the J–V Characteristics of Organic Solar Cells,” *IEEE J. Photovoltaics*, vol. 10, no. 2, pp. 514–521, Mar. 2020, doi: 10.1109/JPHOTOV.2020.2965401.

[34] S. Ono and K. Ohno, “Origin of Charge Transfer Exciton Dissociation in Organic Solar Cells,” in *Excitons*, S. L. Pyshkin, Ed. InTech, 2018. doi: 10.5772/intechopen.69854.

МИКРОТАЛАСНА ТЕХНИКА, ТЕХНОЛОГИЈЕ И СИСТЕМИ
/
MICROWAVE TECHNIQUE, TECHNOLOGIES AND SYSTEMS

(MT/MTI)

Modeling a Planar Circular Loop Antenna using Artificial Neural Networks

Ksenija Pešić[1], Zoran Stanković[2] and Nebojša Dončov[3]

Abstract - **In this paper we present a neural model for planar circular loop antenna based on multilayer perceptron (MLP) network. This model is realized by means of two coupled MLP networks which separately provide the resonant frequency and the minimum value of S_{11} parameter for specified antenna dimensions. The model is trained within the certain range of input parameters (radius of the antenna and the ratio of the radius and the width of loop antenna) that allows for design of planar circular loop antennas operating in the frequency band (500 - 2800) MHz.**

Keywords - **Circular Loop Antenna, Neural model, Neural network**

I. Introduction

Thanks to the development of microwave technology and the introduction of new technologies for conducting electromagnetic (EM) waves based on the use of printed planar structures, realized in microstrip and other similar techniques, printed antennas are being developed [1]. Simplicity production, small dimensions and weight, adaptability to the housings in which they are installed, low price, the ability to work in multiple frequency bands and large mechanical reliability are the features of printed antennas that justify their exceptional popularity and use [1,2]. All these advantages compensate for the disadvantages of printed antennas, such as low bandwidth and low gain, which are a consequence of the manufacturing technology.

Analytical techniques that solve the problems of propagation of EM waves, based on appropriate solutions of Maxwell's equations, are usually applicable for design of antennas with simple configuration. Since in general all antennas, including printed antennas, in real application can be very complex, the use of a numerical model of Maxwell's equations is the only alternative to analytical solutions. There is a large number of numerical techniques able to provide an efficient solution of EM problems, such as the finite element method (FEM) [3], the method of moments (MoM) [3,4], the split-step parabolic equation (SSPE) [4] method, the finite difference time-domain (FDTD) method [3,4] and the transmission line matrix (TLM) method [3,5]. However, these techniques are numerically demanding and careful programming is required in order to reduce time and memory consumption. During antenna design, it is usually necessary to consider a large number of geometrical parameters combinations in order to find a configuration that provides its optimal working performances in a frequency range of interest. Therefore, a process of designing antenna by using any of available numerical techniques can be even more time and memory demanding. Alternative to classical EM simulators may be a use of antenna models based on Artificial Neural Network (ANN) [6,7]. The process of developing a neural model can be also difficult and time consuming as for classical EM models, however when a neural model is successfully developed its speed exceeds the speeds of classical EM models [7-10]. Antenna modeling by using ANN has recently attracted more attention because of the convenience they offer [8-12]. The ANN approach proved to be good in [13,14] when it comes to modeling the resonant frequency and minimum values of S_{11} parameter of planar bowtie dipole and microwave slot antennas considered in [13] and [14], respectively.

In contrast to [13,14] where a model based on one multilayer perceptron (MLP) neural network with two outputs was used, in this paper we propose a neural model that consists of two MLP networks with one output. Each MLP network separately provides one antenna output parameter for input parameters of antenna. In this way, a dependence of each output parameter on input parameters is modeled independently allowing for better accuracy of neural model. The proposed model accuracy and efficiency is illustrated here on the example of design of planar circular loop antenna. Model is able to provide the resonant frequency and the minimum value of S_{11} parameter for specified circular loop antenna dimensions (radius of the antenna R and the ratio of the radius and the width of loop antenna k). Such designed antenna due to its radiation characteristics and the fact that it is soft, portable and very grateful for the extension of the frequency range because it can retain the low value of S_{11} parameter is a good candidate to be used in real-time locating system (RTLS). Nowadays, there is an increasing use of RTLS system that can accurately locate, track and manage assets, inventory or people and help companies make decisions based on the collected location data [15].

II. EM Model of the Planar Circular Loop Antenna

Fig. 1 shows the architecture of a planar circular loop antenna used for both the EM model and the neural model. The basic geometrical parameters are: $R-$ radius of the antenna and $d-$ the width of the loop antenna. Loop antennas are usually classified into two categories, electrically small (circumference is usually less than about one-tenth of a wavelength) and electrically large (circumference is about a free-space wavelength).

The input impedance of a small circular loop antenna made of a thin wire of radius a can be determined by an approximate formula [1]:

Ksenija Pešić[1], Zoran Stanković[2] and Nebojša Dončov[3] are with University of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mails: [ksenija.pesic, zoran.stankovic, nebojsa.doncov]@elfak.ni.ac.rs

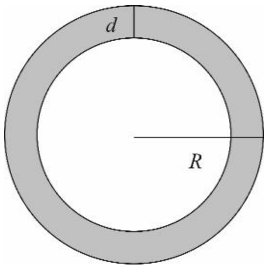


Fig. 1. Architecture of a planar circular loop antenna

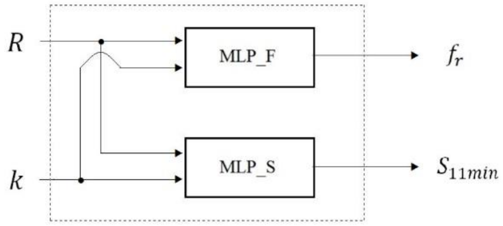


Fig. 2. The neural model of the planar circular loop antenna

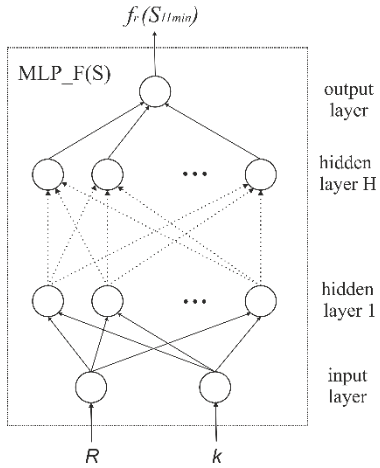


Fig. 3. Architecture of MLP_F(S) neural network

$$Z_{in} = \left(\frac{7.86 \cdot Rf}{10^8}\right)^4 + j \cdot 2\pi f \mu R \left(\ln\left(\frac{8R}{a}\right) - 1.75\right) \quad (1)$$

where f and μ represent the frequency of the signal and the magnetic permeability of the medium, respectively. When a circular loop antenna is made in a planar shape, its cross section should also be taken into account, so analyzes of frequency characteristics of antenna is much more complex and it cannot be performed with a direct usage Eq. (1). This equation will not give accurate results and instead EM simulator has to be used. Using the functions of the Toolbox Antenna in the MATLAB software environment [16], an EM model can be created in order to analyze the characteristics of this antenna. Antenna Toolbox uses the method of moments (MoM) to compute antenna properties such as input impedance, surface properties such as current and charge distribution, and field properties such as the near-field and far-field radiation pattern [16,17]. The basic function in Antenna Toolbox for creation of the circular antenna object is *loopCircular(Name,Value)*. This function creates a one wavelength circular loop antenna, with additional properties specified by one, or more name-value pair arguments [16]. Due to the small size of the antenna, the ratio of radius to width was used, and not the width of the antenna itself. This ratio is defined as $k = R / d$, where R is the radius of the loop and d is the width of the antenna. Accordingly, in the EM model, the variable parameters are the radius of the antenna and the previously defined ratio k. On the created loop antenna object we can use standard functions for network configuration in MoM method, antenna structure presentation, radiation pattern presentation, estimation of radiation in the plane of azimuth and elevation (*show*, *pattern*, *patternAzimuth*, *patternElevation*, *sparameters*) [16,17].

III. Neural Model of the Planar Circular Loop Antenna

The neural model of the planar circular loop antenna consists of two multilayer perceptron networks (MLP_F and MLP_S) and its architecture is shown in Fig.2. The main purpose of the model is to perform the mapping from a space of physical parameters of the antenna (radius of the antenna R and the ratio of the radius and the width of loop antenna k) to a space of antenna operating characteristics consisting of resonant frequency f_r

$$[f_r] = f_{MLP\_F}(R, k) \quad (2)$$

and minimum value of S_{11} parameter (value of S_{11} parameters at antenna resonant frequency - S_{11min})

$$[S_{11min}] = f_{MLP\_S}(R, k) \quad (3)$$

at the given feed line impedance (z_f =const). Neural networks MLP_S and MLP_F have the same architecture which is represented in Fig. 3.

In matrix representation, the neural model will have the vector of input variables $\mathbf{x} = [R \quad k]^T$. The vector of output variable in the MLP_F network will be $\mathbf{y} = [f_r]$, while in the MLP_S network the vector of the output variable is $\mathbf{y} = [S_{11min}]$. MLP neural network for both cases can be described by $\mathbf{y} = y(\mathbf{x}, W, B)$, where y is a network processing function, W is a set of connection weighting matrices $\mathbf{w}_i$, $W = \{\mathbf{w}_1, \mathbf{w}_2, \ldots \mathbf{w}_{H+1}\}$ and B is a set of bias vectors $\mathbf{b}_i$, $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots \mathbf{b}_{H+1}\}$ (H is the total number of hidden layers of the MLP network). Accordingly, part of the neural model for the MLP_F and MLP_S networks can be described as:

$$[f_r] = y([R \;\; k]^T, W, B) \quad (4)$$

$$[S_{11min}] = y([R \;\; k]^T, W, B) \quad (5)$$

respectively. Output of MLP l-th hidden layer, $\mathbf{y}_l$, is represented by the following function:

$$\mathbf{y}_l = F(\mathbf{w}_l \mathbf{y}_{l-1} + \mathbf{b}_l) \quad l = 1,2, \ldots H \tag{6}$$

where $\mathbf{y}_{l-1}$ vector represents the output of (l-1)-th hidden layer, the output of the input layer is a vector $\mathbf{y}_0 = \mathbf{x}$, $\mathbf{w}_l$ is a connection weight matrix among (l-1)-th and l-th hidden layer neurons and $\mathbf{b}_l$ is the vector containing biases of l-th hidden layer neurons. Hyperbolic tangent sigmoid transfer function:

$$F(u) = \frac{e^u - e^{-u}}{e^u + e^{-u}} \tag{7}$$

is used as an activation function of neurons in hidden layers. Activation function of neuron in the output layer is linear so that the output of MLP network is:

$$\mathbf{y} = \mathbf{w}_{H+1} \mathbf{y}_H + \mathbf{b}_{H+1} \tag{8}$$

where $\mathbf{w}_{H+1}$ is the connection weighing matrix between neurons of H-th hidden layer and neurons of output layer and $\mathbf{b}_{H+1}$ is the vector containing biases of output layer.

The notation for such defined MLP neural model, that will be used further in the paper, is MLPH-N_1-…-N_i-…-N_H. N_i represents the number of neurons in the i-th hidden layer. Each network was trained three times with new initial connection weights and thresholds, whose values are random numbers in the interval [-1 1], in order to obtain the best trained model. For example, the notation MLP2-22-15 is used for the MLP model whose neural network has a total of 4 layers (input, output and two hidden layers), this model has 22 neurons in the first and 15 neurons in the second hidden layer.

Neural model was implemented in the MATLAB software environment [16]. For MPL training and testing we generated training and test sample sets by using the EM model of the circular loop antenna described in section II. This data sets have the following format $\{(\mathbf{x}^t, \mathbf{y}^t)\}$, or format $\{([R^t\ k^t], [f_r^t\ S_{11min}^t])\}$ where $\mathbf{x}^t$ is the vector of input combinations of the variables R^t and k^t, while $\mathbf{y}^t$ is vector that contains desired outputs of the neural network f_r^t and S_{11min}^t for given input. The notation t in the superscript means that the samples are belonging to training and testing set. The range of input parameters for which the network is trained is $R\ [m] \in\ [0.020 - 0.100]$ and $k \in [10 - 100]$. This range of input parameters provides achieving resonant frequency for which holds $f_r^t = [500 - 2800]\ [MHz]$, that means that the output band of the model belongs to UHF band as defined according International Telecommunication Union (ITU).

By using the EM model of the planar circular loop antenna and uniform distribution of samples, the following sets for training and testing were generated:

$$\left\{([R^t\ k^t]^T, [f_r^t,\ S_{11\min}^t]) \mid R^t \in [0.02 : d_1 : 0.100], k^t \in [10 : d_2 : 100]\right\} \tag{9}$$

where parameters d_1 and d_2 represent steps and their values d_1 = 0.005 m and d_2 = 2 were used to generate the training set, the values d_1 = 0.007 m and d_2 = 3 were used for the test set. As a result, the training and test sets with 782 and 372 samples, respectively, were obtained. Levenberg-Marquardt training method [7] was used during the model training. Since there are two MLP networks (MLP_F and MLP_S), the appropriate columns from the test and training files are used for each of the networks. In order to obtain the best possible model, the training of a number of MLP neural networks with two hidden layers and different number of neurons in them, was performed. In this process, training and test sets were preprocessed by normalizing the inputs and targets so that they were in the interval [-1,1]. During the training, the target value of the mean square error of the model outputs on the training set was 10^{-4}. During the testing of both networks, three test metrics were observed: values of the worst case error (WCE), values of average test error (ATE) and correlation coefficient [7] for both network outputs, $r^{PPM}(f_r)$ and $r^{PPM}(S_{11\min})$. The goal was to find a neural network with good generalization properties. The testing results for the six MLP_F neural networks with the best test statistics are shown in Table I.

TABLE I
TESTING RESULTS FOR SIX MLP_F NETWORKS WITH THE BEST TEST STATISTICS

MLP_F	WCE[%]	ATE[%]	$r^{PPM}(f_r)$
MLP2-20-15	0.4702	0.1043	0.99999003
MLP2-12-8	0.5881	0.1399	0.99998983
MLP2-17-9	1.1560	0.1775	0.99998961
MLP2-9-6	0.7387	0.1962	0.99998829
MLP2-9-5	0.6632	0.1334	0.99998768
MLP2-10-8	1.1873	0.2896	0.99998756

For the MLP_F part of the neural model, the neural network MLP2-20-15 was chosen, which has the highest value of the correlation coefficient and the lowest value of the worst case error (WCE). The scattering diagram of the MLP2-20-15 neural network on the test set for the output f_r is shown in Fig. 4. The output of the neural model f_r has a very little scatter in relation to the reference values.

The testing results for the six MLP_S neural networks with the best test statistics are shown in Table II.

TABLE II
TESTING RESULTS FOR SIX MLP_S NETWORKS WITH THE BEST TEST STATISTICS

MLP_S	WCE[%]	ATE[%]	$r^{PPM}(S_{11\min})$
MLP2-5-5	26.6089	2.8374	0.9759
MLP2-6-6	28.1494	2.8188	0.9754
MLP2-9-6	25.4494	2.8850	0.9734
MLP2-8-8	26.3811	2.9488	0.9731
MLP2-10-10	24.0547	2.9181	0.9730
MLP2-10-4	30.1837	2.8648	0.9715

For the MLP_S part of the neural model, the neural network MLP2-5-5 was chosen, which has the highest value of the correlation coefficient. The scattering diagram of the MLP2-5-5 neural network on the test set for the output S_{11min} is shown in Fig. 5.

In Fig.5 it can be seen that for $S_{11min} \geq$ -20 dB the scattering for this output of the model is small to moderate. For $S_{11min} <$ -20 dB the scattering is more pronounced which is expected due to the presence of sharp deep peak change in S_{11} values around the resonant frequency.

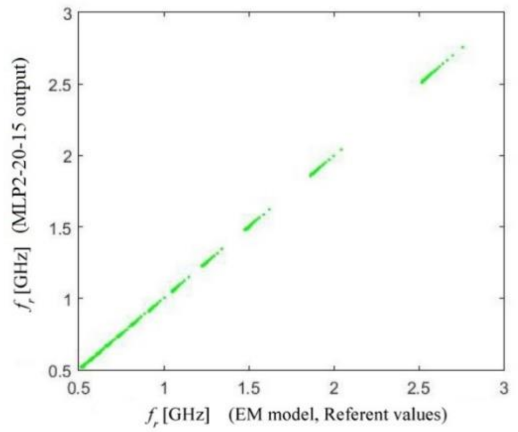


Fig. 4. Scattering diagram of MLP2-20-15 neural model on the test set for f_r output

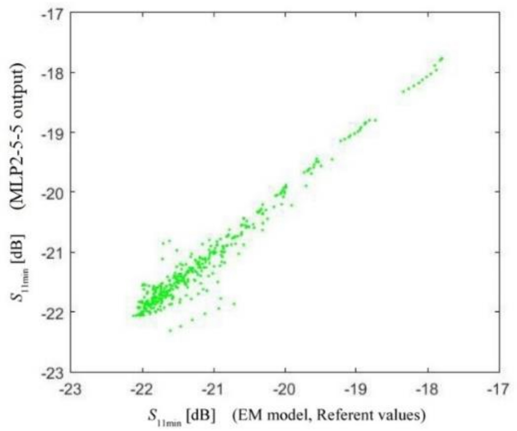


Fig. 5. Scattering diagrams of MLP2-5-5 neural model on the test set for S_{11min} output

However, in that case antenna is already well matched so the reduced model accuracy will not significantly limit its use for an analysis of antenna frequency characteristics such as determining the boundaries and bandwidth of antenna operating frequency range.

Time required to the proposed neural model to simulate 372 points of the test set was 0.04 seconds, while for the same number of points the EM model performed the simulation in 3.23 seconds. Both simulations were performed on a platform whose processor is Intel Core i7 and RAM 12 GB.

IV. Conclusion

Using neural models of planar circular loop antenna, results can be obtained faster and in a simpler way by avoiding complex calculations of Maxwell’s equations that take a very long time. The proposed neural model has shown satisfactory accuracy which is very close to the accuracy of EM simulator. Further development will be focused on the optimization of this model, in order to provide the possibility of quickly finding the input parameters for the desired output quantities.

V. Acknowledgment

This work was supported by the Ministry of Education, Science and Technological Development of Republic of Serbia (Grant No. 451-03-9/2021-14/200102).

References

[1] Balanis, C. A., *Antenna Theory: Analysis and Design*, Wiley, New York, 2005.
[2] R. Waterhouse, *Printed Antennas for Wireless Communications*, John Wiley & Sons, Ltd, 2007.
[3] Matthew N. O. Sadiku, *Numerical techniques in electromagnetics*, 2nd edition, CRC Press LLC, 2001.
[4] Levent Sevgi, *Electromagnetic Modeling and Simulation*, First Edition. The Institute of Electrical and Electronics Engineers, Inc. Published by John Wiley & Sons, Inc., 2014
[5] J. Joković, T. Dimitrijević, N. Dončov, “Efficient Modeling of a Circular Patch-Ring Antenna Using the Cylindrical TLM Approach”, *IEEE Antennas and Wireless Propagation Letters*, Vol. 16, No. 4, pp. 2070-2073, 2017.
[6] S. Haykin, *Neural Networks*, New York, IEEE, 1994.
[7] Q. J. Zhang, K. C. Gupta, *Neural networks for RF and microwave design*, Boston, MA: Artech House, 2000.
[8] M. Milijić, Z. Stanković and B. Milovanović, "Efficient model of slotted patch antenna based on neural networks,"*2009 9th International Conference on Telecommunication in Modern Satellite, Cable, and Broadcasting Services*, Nis, 2009, pp. 384-387.
[9] M. Pandit, T. Bose, “Application of Neural Network Model for designing Circular Monopole Antenna”, *International Symposium on Devices MEMS, Intelligent Systems & Communication (ISDMISC) - Proceedings published by International Journal of Computer Applications (IJCA)*, April 2011, Gangtok, Sikkim, India, pp. 18 -21.
[10] M. Chetioui, A. Boudkhil, N. Benabdallah and N. Benahmed, "Design and optimization of SIW patch antenna for Ku band applications using ANN algorithms,"*2018 4th International Conference on Optimization and Applications (ICOA)*, Mohammedia, 2018, pp. 1-4. DOI: 10.1109/ICOA.2018.8370530.
[11] U. Ozkayaa, L. Seyfia, “Dimension Optimization of Microstrip Patch Antenna in X/Ku Band via Artificial Neural Network”, *Procedia - Social and Behavioral Sciences* 195 (2015), pp. 2520 – 2526.
[12] R. Ghayoula, A. Gharsallah, N. Fadlallah and M. Rammal, "Artificial Neural Networks ANN Approach for the Synthesis of Patch Antennas,"*2007 IEEE Radar Conference*, Boston, MA, 2007, pp. 285-290. DOI: 10.1109/RADAR.2007.374229
[13] M. Sarevska, Z. Stankovic, N. Doncov, I. Milovanovic and B. Milovanovic, "Design of well-matched UHF Planar Bowtie Dipole Antenna using Neural Model," 2019 14th International Conference on Advanced Technologies, Systems and Services in Telecommunications (TELSIKS), Nis, Serbia, 2019, pp. 331-334, doi: 10.1109/TELSIKS46999.2019.9002319.
[14] Z. Stankovic, N. Doncov, B.Stosic, M. Sarevska and I. Milovanovic, "Design of well-matched Microwave Slot Antenna on a Flat Metal Grounded Plate using Neural Model", 2020 Information, Communication and Energy Systems and Technologies (ICEST)
[15] https://www.atlasrfidstore.com/what-is-rtls-an-introduction-to-real-time-location-systems/
[16] https://www.mathworks.com
[17] http://ece.wpi.edu/mom/

Modelling of Conformal Antennas using Time-Domain TLM Method

Tijana Dimitrijevic, *Member, IEEE*, Ekrem Altinozen, *Student Member, IEEE*, Aleksandar Atanaskovic, *Member, IEEE*, Jugoslav Jokovic, *Member, IEEE*, Ana Vukovic, *Member, IEEE*, Phillip Sewell, *Senior Member, IEEE*, and Nebojsa Doncov, *Member, IEEE*

***Abstract*—As textile antennas are seeing a major growth in recent years, there is a demand for understanding how they behave when exposed to a realistic environment. Typically, that is done by analyzing the effect of bending on antenna performance. To achieve that, an appropriate, accurate and reliable modeling approach is required. This paper uses two computational methods that employ different discretisation approaches to analyse the effect of bending on antenna performance, namely antenna reflection coefficient. The paper further aims to investigate how the discretisation approach influences the nature and the accuracy of results.**

***Index Terms*—Bending effect, textile antennas, patch antennas, TLM method.**

I. Introduction

WEARABLE applications have experienced an enormous rise in a recent decade fueled by rapid development of 5G technology and continued demand for better health and wellbeing. One of the current challenges in designing wearable antennas is to account for realistic environment which may include the effect of the human tissue but also the impact of arbitrary deformations of textile antenna caused by human movement in everyday activities. Some initial results on the impact of human tissue have been reported in [1-3], showing that the level of performance degradation is dependent on the antenna design.

For the case of arbitrary antenna deformation, a focus has been placed on the special case of cylindrical bending which is also the simplest way of deformation. An analytical model of the cylindrical bending using the cavity mode theory has been presented in [4] with the conclusion that conformally mounted patch antenna that has substrate thickness less than one tenth of the bending radius experiences less than 5% change in resonant frequency and can be therefore treated as a planar antenna [4]. A number of papers used either experimental or simulation approaches to evaluate impact of bending on antenna performance [5-10] and show that bending of antenna can affect the resonant frequency of the antenna. Antenna can be bent along the length of the patch (E-plane) or along the width of the patch (H-plane). Results have shown that E-plane bending has a more significant impact on antenna performance as it directly affects the length of antenna which in turn determines the resonant frequency [7-8]. More recently it has been shown that the thickness of the substrate also plays an important role on how antenna will experience the bending deformation [10].

Whilst experimental measurements are necessary for full deployment, the electromagnetic (EM) computer analysis and simulation plays a vital role in the overall design process. It is important to note that the case of cylindrical bending is the one that can be described using Boolean geometry and as such can be analyzed using commercial time domain EM software that typically uses cubic grid for discretization of the problem space. However, the geometry of the cylindrically bent antenna does not conform to cubic mesh and in order to develop an accurate numerical model the antenna geometry needs to be sampled using a very fine mesh which in turn can demand large computation resources. For example, for time domain computational methods, very fine cubic mesh results in a very small timestep and long runtime. On the other hand, discretization mesh such as cylindrical or tetrahedral mesh that conforms to the structure is preferred as it eliminates the discretization errors and results in much faster and more accurate simulations.

In this paper, the well-established, time domain numerical TLM method has been used to analyze the impact of the bending on textile patch antenna on the S_{11} parameter. The paper considers only the case of E-plane bending as it has a more severe impact on antenna performance. The paper compares the antenna results for several bending angles obtained using TLM method based on the rectangular mesh (recTLM) [11], and the TLM method based on purely cylindrical mesh (cylTLM) [11]. Computational requirements needed for converged results are also compared between these two methods. The paper also seeks to understand whether the type of the mesh used would influence the nature of results. In the following sections a brief description of the TLM method is given, followed by the results and conclusion section.

Tijana Dimitrijevic, Aleksandar Atanaskovic, Jugoslav Jokovic and Nebojsa Doncov are with the Faculty of Electrical Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia (e-mail: tijana.dimitrijevic@elfak.ni.ac.rs, aleksandar.atanaskovic@elfak.ni.ac.rs, jugoslav.jokovic@elfak.ni.ac.rs, nebojsa.doncov@elfak.ni.ac.rs).

Ekrem Altinozen, Ana Vukovic and Phillip Sewell are with the Faculty of Engineering, University of Nottingham, Nottingham, UK, (e-mail: ekrem.altinozen1@nottingham.ac.uk, ana.vukovic@nottingham.ac.uk, phillip.sewell@nottingham.ac.uk).

II. The TLM Methodology

The TLM method is a time-domain numerical modeling technique based on equivalences between Maxwell's equations and equations for voltage and current pulses propagation along transmission lines to determine electromagnetic field components [11]. It uses a network of interconnected nodes filling out the propagation space while appropriately representing electromagnetic properties of the homogeneous or inhomogeneous media. A basic TLM cell is the symmetrical condensed node (SCN), but usually HSN (hybrid SCN) node is used to enable modelling of a non-uniform mesh [11]. Depending on the modeling problem, nodes can be generated as cubes or cuboids for Cartesian mesh, "slice of cake" type node for a cylindrical mesh [11] or tetrahedral nodes for an unstructured mesh [12]. Different types of nodes, a cuboid in a Cartesian grid, a part of a slice of cake in a cylindrical grid, and a tetrahedral one, are illustrated in Fig.1.

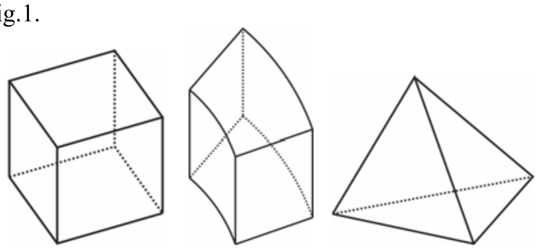


Fig. 1. TLM cell: a) rectangular, b) cylindrical, c) tetrahedral.

In the conventional TLM algorithm, in both rectangular and cylindrical grid, a problem geometry and initial conditions are firstly defined, followed by the calculation of equivalent voltages and currents for each cell, which may be further used to calculate the desired electromagnetic field component. The main algorithm consists of two basic processes that are iteratively repeated for all the nodes within the computational area, and these are, scattering, where reflected voltages at each node are calculated from the incident voltages, and connection where reflected pulses become incident pulses to the neighboring nodes for the next time step. The simulation starts by defining excitation voltages which then propagate through series of reflections and scattering between adjacent nodes while different propagation conditions pertaining to pulse velocity and boundaries are considered. Hence, a dielectric presence is characterized by its relative permittivity and loss tangent. As an output of the simulation, a voltage or a current induced in the wire can be obtained which is further manipulated to determine the reflection coefficient [13].

III. Results and Discussion

The textile patch antenna bent over a cylinder is schematically shown in Fig.2. The patch has the length l and the width w and is bent over the cylinder of radius R over angle 2θ. A coaxial feed is modelled as a wire which is used as an excitation. For the patch antenna considered in this paper, both the patch and the ground plane are described as PEC layers, while the wire is modelled using the compact wire model adapted to the cylindrical mesh [14]. The length and the width of the patch are $l = 50$ mm and $w = 39.5$ mm, respectively, while the length and the width of the substrate and ground plane are $W = L = 100$ mm. It is realized on the substrate of the relative permittivity 2.1 and the thickness $h = 2$ mm. A coaxial feed is placed at 11.5 mm from the patch edge. The flat antenna is designed to resonate at 2.45 GHz.

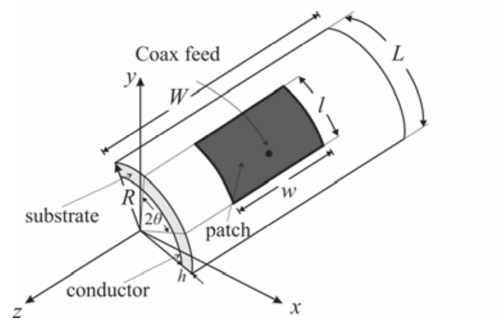


Fig.2. A rectangular patch antenna with a bending 2θ in an E-plane.

To investigate an influence of a curvature on the antenna performances several models of the antenna are considered, and each model is generated for a specific bending angle whilst preserving patch dimensions. All models are meshed with the rectangular and cylindrical TLM mesh. A convergence analysis has been conducted for both recTLM and cylTLM to determine the most adequate mesh. The convergence analysis of the recTLM method is done for the bent antenna with $2\theta = 25$ degrees and for cylindrical case the convergence analysis is done for the case of flat antenna. In both cases the methods consider the structures that do not conform to a particular mesh type. In the case of cylTLM, the flat antenna has been designed by applying very small bending angle, i.e., $2\theta = 0.1$ degrees (corresponding to the radius of $R = 22.63$ m in a cylindrical grid), i.e. $R>>l$, hence it can be considered as the flat one. Fig.3 shows the convergence of the resonant frequency for different mesh size Δl, where Δl represents the cell size within an area around the structure with a refined mesh applied. In the cylTLM method, this area is simply the substrate area.

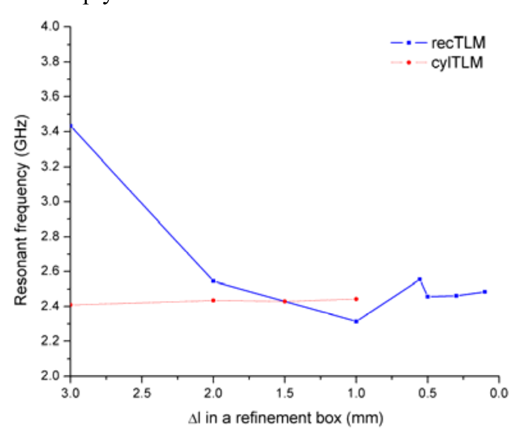


Fig.3. Illustration of the resonant frequency convergence with the mesh refinement when the recTLM method (blue line) is used for the bent antenna, and cylTLM method (red line) is used for the flat antenna (Δl represents the observed cell size).

As can be seen, the resonant frequency of the antenna converges to a different value in the recTLM and cylTLM method. This is as expected since the bending affects the resonant frequency. Furthermore, the recTLM achieves convergence for much smaller values of Δl confirming that discretization error around the curved surfaces can affect the convergence analysis. The cylTLM on the other hand has much smoother and faster convergence since the mesh is ideally suited to the structure even though the considered antenna is almost flat resulting in 10 times smaller mesh for the converged result compared to recTLM.

The comparison of a number of cells and their dimensions required for different mesh types is shown in Table 1. It shows that rectangular TLM requires much larger computational resources than cylindrical TLM, resulting in about 6 times larger mesh when 0.5 mm rectangular cell size is used, and about 360 times larger mesh for 0.1 mm cell size. This is a significant advantage of cylindrical TLM method for this particular application.

Table 1. The comparison of computational mesh size between recTLM and cylTLM approach

Method	Cell size (general mesh/ refined area)	Number of nodes	Wire radius
recTLM	1 mm/0.5 mm	80×280×260 ~6M	0.1mm
recTLM	1 mm/0.1 mm	240×1240×1140 339M	0.025mm
cylTLM	1.449 mm/1 mm	151×42×149 944k	0.1mm

For the case of flat antenna, the cylTLM method with 1 mm mesh gives resonant frequency at 2.443 GHz and the recTLM gives resonant frequency at 2.454 GHz if 0.5 mm mesh is used and 2.492 GHz if 0.1 mm mesh is used. Further investigation has included modeling of antennas cases with a different bending angle, i.e., 2θ=25, 50 and 60 degrees, while patch dimensions are kept the same. The S_{11} results obtained using cylTLM, and recTLM methods are shown in Figs. 4 and 5, respectively. The S_{11} of the flat antenna is also included in figures for reference. All results show that resonant frequency values are influenced by the curvature. According to the Fig.4, when the cylTLM approach is used, results show that the resonant frequency is increased with rising the bending angle in an almost linear fashion, while the matching condition is affected as well. However, in the case of recTLM method, as presented in Fig.5a, the trend is different for coarser mesh (0.5 mm), showing reduction in resonant frequency for various bending angles compared to Fig.4. When a finer mesh is used (0.1mm), as shown in Fig.5b the trend is more similar to the cylTLM results. A difference between the resonant frequency values by two approaches might be attributed to introducing the wire for excitation of much smaller wire radius in the recTLM than in the cylTLM approach, due to a much smaller cell size. Also, a better matching is possible to achieve when a finer rectangular mesh is used.

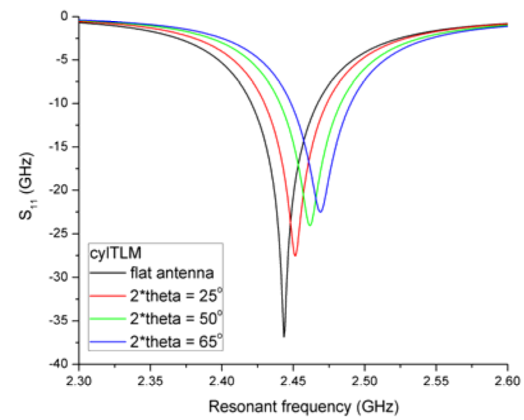
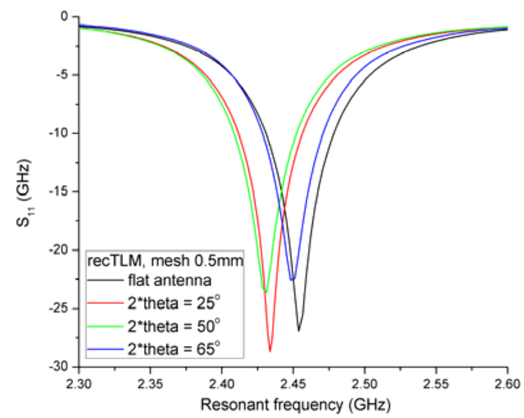
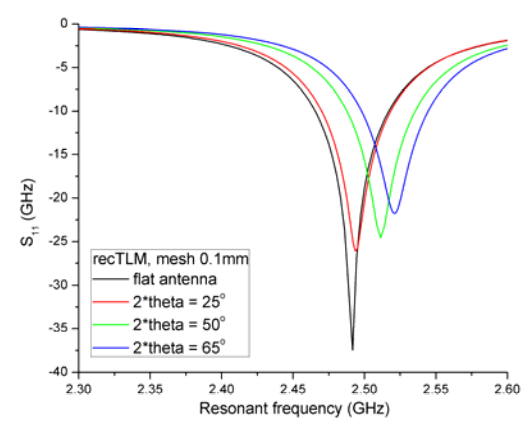


Fig.4. Comparison of S11 parameter of the flat antenna and antenna bent over a cylinder of angle 2θ =25, 50 and 60 degrees. Results are obtained using cylTLM method (1.0 mm cell size).



a)



b)

Fig.5. Comparison of S11 parameter of the flat antenna and antenna bent over a cylinder of angle 2θ =25, 50 and 60 degrees. Results are obtained using recTLM method with a) courser mesh (0.5 mm cell size), b) finer mesh (0.1 mm cell size).

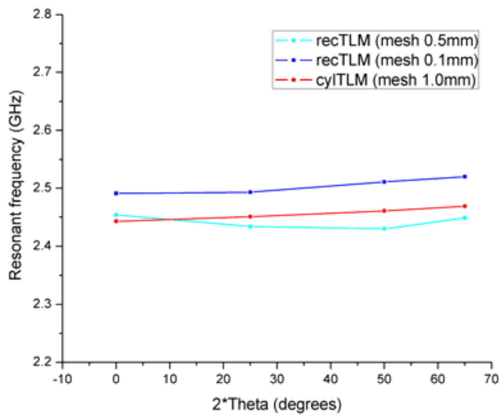


Fig.6. The resonant frequency vs the bending angle (2*theta) of the flat and bent rectangular patch antenna obtained by recTLM, and cylTLM approaches.

Fig. 6. presents the comparison of the resonant frequency reached via these two methods for various bending angles. Fig.7. reveals that the maximum frequency shift when the cylTLM approach is applied is 26 MHz, while the recTLM approach with the courser mesh gives 13 MHz, and 29 MHz for the finer mesh.

According to the presented results and the ease of convergence of the cylTLM method it can be concluded that results obtained using cylTLM method agree much better with the reported literature stating that E-plane bending causes the resonant frequency to shift to higher frequencies. The results obtained using the recTLM method are heavily affected by the discretization error which is confirmed by comparing results obtained using 0.5 mm and 0.1 mm mesh. It can be concluded that only a very fine mesh is needed so that the discretization error is not affecting the main EM phenomena.

IV. Conclusion

In this paper, the well-established TLM method is applied to investigate the influence of the curvature on the rectangular patch antenna resonant frequency. Two different types of the TLM mesh, based on different cell geometries are used, namely the rectangular, and the cylindrical TLM method. Results show that when using rectangular TLM method a very fine mesh must be used to minimize the discretization error and observe correct behavior. This is not the case for the cylTLM method as the mesh is perfectly aligned with the structure resulting in a tenfold increase in the mesh size and significant reduction in computation time and memory. The converged results of both methods show that when antenna is bent in the E-plane the resonant frequency of antenna shifts to higher frequencies.

Acknowledgment

This work was supported by the Ministry of Education, Science and Technological Development of Republic of Serbia (Grant No. 451-03-9/2021-14/200102), Science Fund of the Republic of Serbia (Grant No. 6394135), and the Royal Society International Exchanges Grant IES\R1\201311.

References

[1] J. S. Colburn and Y. Rahmat-Samii, "Human proximity effects on circular polarized handset antennas in personal satellite communications," in *IEEE Transactions on Antennas and Propagation*, vol. 46, no. 6, pp. 813-820, June 1998, doi: 10.1109/8.686767.

[2] S. S. Zhekov and G. F. Pedersen, "Effect of Dielectric Properties of Human Hand Tissue on Mobile Terminal Antenna Performance," *2020 14th European Conference on Antennas and Propagation (EuCAP)*, 2020, pp. 1-5, doi: 10.23919/EuCAP48036.2020.9135910.

[3] M. Koohestani, N. Pires, A. K. Skrivervik and A. A. Moreira, "Performance Study of a UWB Antenna in Proximity to a Human Arm," in *IEEE Antennas and Wireless Propagation Letters*, vol. 12, pp. 555-558, 2013, doi: 10.1109/LAWP.2013.2259212.

[4] C. Krowne, "Cylindrical-Rectangular Microstrip Antenna. Antennas and Propagation," *IEEE Transactions on. AP*-31, 194 – 199, 1983.

[5] I. Locher, M. Klemm, T. Kirstein and G. Troster, "Design and Characterization of Purely Textile Patch Antennas," in *IEEE Transactions on Advanced Packaging*, vol. 29, no. 4, pp. 777-788, Nov. 2006, doi: 10.1109/TADVP.2006.884780.

[6] C. Hertleer, A. Tronquo, H. Rogier, L. Vallozzi and L. Van Langenhove, "Aperture-Coupled Patch Antenna for Integration Into Wearable Textile Systems," in *IEEE Antennas and Wireless Propagation Letters*, vol. 6, pp. 392-395, 2007, doi: 10.1109/LAWP.2007.903498.

[7] M. Tanaka and Jae-Hyeuk Jang, "Wearable microstrip antenna," *IEEE Antennas and Propagation Society International Symposium. Digest. Held in conjunction with: USNC/CNC/URSI North American Radio Sci. Meeting (Cat. No.03CH37450)*, 2003, pp. 704-707 vol.2, doi: 10.1109/APS.2003.1219333.

[8] F. Boeykens, L. Vallozzi, H. Rogier, "Cylindrical Bending of Deformable Textile Rectangular Patch Antennas", *International Journal of Antennas and Propagation*, vol. 2012, Article ID 170420, 11 pages, 2012.

[9] M. E. Bakkali, M. Martinez-Estrada, R. Fernandez-Garcia, I. Gil and O. E. Mrabet, "Effect of Bending on a Textile UHF-RFID Tag Antenna," *2020 14th European Conference on Antennas and Propagation (EuCAP)*, 2020, pp. 1-5, doi: 10.23919/EuCAP48036.2020.9135331

[10] L. Song and Y. Rahmat-Samii, "A Systematic Investigation of Rectangular Patch Antenna Bending Effects for Wearable Applications," in *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 5, pp. 2219-2228, May 2018, doi: 10.1109/TAP.2018.2809469.

[11] C. Christopoulos, The Transmission-Line Modeling Method: TLM. New York, NY, USA: Wiley, 1995.

[12] P. Sewell, T. M. Benson, C. Christopoulos, D. W. P. Thomas, A. Vukovic and J. G. Wykes, "Transmission-line modeling (TLM) based upon unstructured tetrahedral meshes," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 6, pp. 1919-1928, June 2005, doi: 10.1109/TMTT.2005.848078.

[13] T. Z. Dimitrijevic, J. J. Jokovic and N. S. Doncov, "Efficient Modeling of a Circular Patch-Ring Antenna Using the Cylindrical TLM Approach," in *IEEE Antennas and Wireless Propagation Letters*, vol. 16, pp. 2070-2073, 2017, doi: 10.1109/LAWP.2017.2696122.

[14] T. Ž. Dimitrijević, J. J. Joković, B. D. Milovanović, N. S. Dončov, "TLM Modelling of a Probe-coupled Cylindrical Cavity based on Compact Wire Model in the Cylindrical Mesh", International Journal of RF and Microwave Computer-Aided Engineering, John Wiley & Sons Inc., Wiley InterScience Periodicals, Vol.22, No. 2, pp. 184-192, 2012. DOI: 10.1002/mmce.20569

Reduced Dimensions Planar Rat Race Coupler Design

Denis A. Letavin and Dusan A. Nesic, *Member, IEEE*

***Abstract*—The design of a rat race directional coupler was investigated in the Cadence AWR Design Environment program. By using low-pass filters instead of quarter-wave sections, it was possible to reduce the size of the device by 82.3%. In this case, the following deterioration of frequency characteristics occurred: narrowing of the operating frequency band by 19.3%, an increase in imbalance, and a decrease in matching. Also, the area of the compact double ring coupler was reduced by 84.5% while the bandwidth was narrowed by 29.2%.**

***Index Terms*—Filter, miniaturization, stub, coupler.**

I. Introduction

THE rat race directional coupler functions as a signal power divider with two phase differences, 0 and 180 degrees, depending on which port is considered to be the input port. The topology of such a device consists of four segments, three of which have an electrical length of 90 degrees, and the remaining 270 degrees. After combining such segments, a ring with a length of 1.5 wavelengths per line is obtained. Due to the fact that the wavelength and frequency are related, the lower the operating frequency of the device, the more area on the printed circuit board it will have. Therefore, at low frequencies, where the couplers turn out to be cumbersome, its miniaturization is actual, with the minimum possible deterioration of frequency characteristics. There are a lot of options for miniaturizing the spokesmen. Let's take a look at just a few. The following approaches are used to miniaturize tap-off devices: microstrip cells [1], T-shaped structures [2], dual-transmission lines [3], using C-SCMRC resonators with distributive equivalent circuit [4], artificial transmission lines [5, 10, 14, 15], resonators [6, 7, 13], circular defected ground structure [8], shunt-stub-based artificial transmission lines [9], multilayer LTCC [11, 12]. In the proposed work, a method is considered for obtaining a compact ring coupler by replacing all quarter-wave sections with low-pass filters. Also considered is a compact double ring coupler, which, due to the addition of one more circuit, increases the operating frequency band and dimensions.

Dusan Nesic is with Centre of Microelectronic Technologies, Institute of Chemistry, Technology and Metallurgy, University of Belgrade, Njegoševa 12, 11000 Belgrade, Serbia (mail: nesicad@nanosys.ihtm.bg.ac.rs)

Denis A. Letavin is with the Institute of Radioelectronics and Informational Technologies, Ural Federal University, Mira str. 32, 620002 Ekaterinburg, Russia (e-mail: d.a.letavin@urfu.ru)

II. Design Coupler

A circular directional coupler is a type of power divider used in microwave technology. Using the Cadence AWR DE program and the built-in TXLine calculator, the topology of a standard coupler was calculated (Fig. 1). The well-known FR4 material acts as a substrate material, with $\varepsilon = 4.4$ and $h = 1$ mm, and as a central frequency of 1 GHz. The obtained frequency characteristics of the coupler in the AWR program are illustrated in Figures 2 and 3. The area of the device is 43 x 91.5 = 3934.5 mm^2.

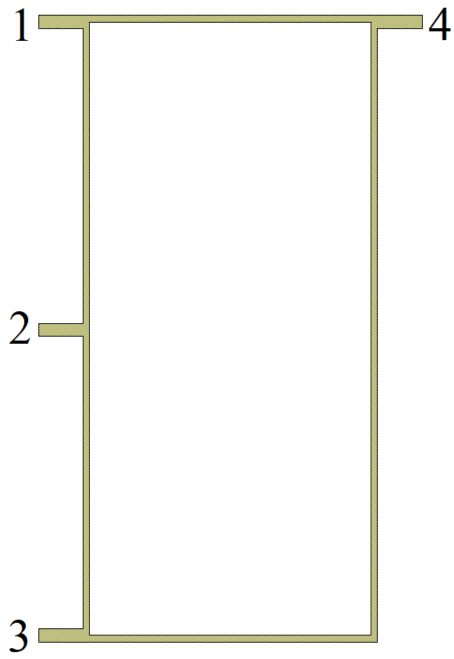


Fig. 1. Standard rat race coupler layout

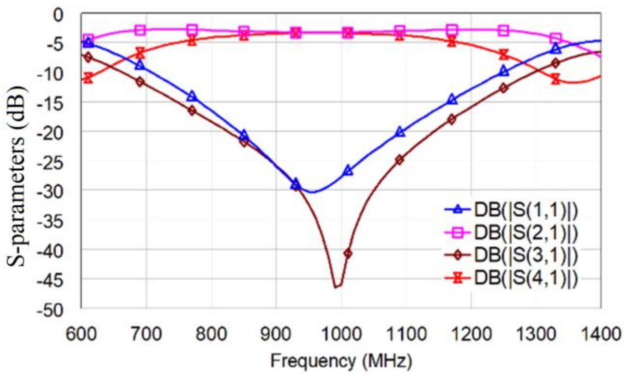


Fig. 2. S-parameter versus frequency plot for a standard rat race coupler

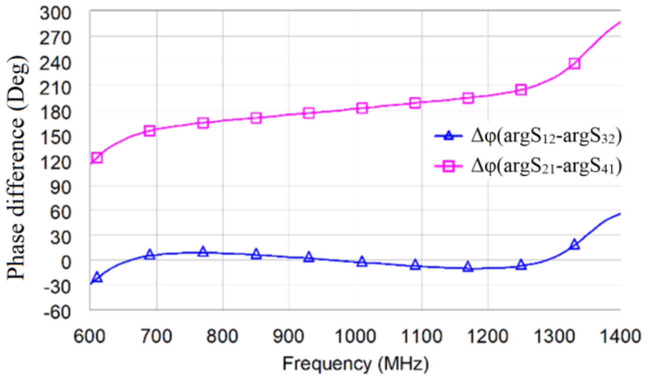


Fig. 3. The graph of the phase difference of the output signals depending on the choice of the input port (blue line - 1 input port, pink line - 4 input port)

According to the graphs obtained, it can be seen that the operating frequency band of the coupler at the isolation level "minus" 20 dB is 320 MHz (824-1144 MHz). The imbalance between the output signals of the structure in the frequency band does not exceed 1.3 dB. The area inside the coupler is not used in any way, which is an additional disadvantage of such a device. To reduce the size of the ring coupler, the required low-pass filters are initially calculated, which are used instead of the quarter-wave sections. For this purpose, a filter is calculated for the required wave impedance of 70.7 Ohm, which has a phase incursion of 90 degrees at a central frequency of 1 GHz. Filters can be implemented in a T-shaped or U-shaped circuit (Figure 4). They are completely equivalent to each other and can be used with equal efficiency. However, in our case, a U-shaped circuit was chosen, due to the fact that the extreme capacitive elements can be combined with the extreme elements from other filters, which can make it possible to more successfully fill the space inside the coupler. Low-pass filters can be calculated using the built-in iFilter tool in the Cadence AWR software.

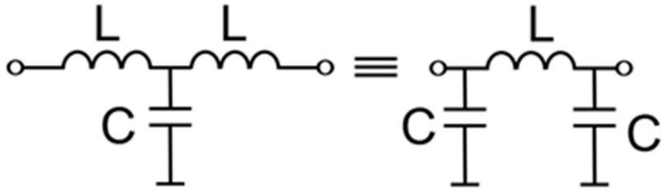


Fig. 4. T-shaped and U-shaped low-pass filter circuit

To calculate the ratings of CLC (capacitance - inductance - capacitance) elements, it is necessary to know the characteristic impedance of the transmission line, the parameters of the substrate material, the central frequency at which it is necessary to provide a phase shift of 90 degrees. After all filters are calculated, they are gradually installed instead of the quarter-wave sections of the coupler. The arrangement of the filter elements is carried out in such a way that there is no electrical contact between the elements of adjacent filters. After this procedure, an electrodynamic design calculation is performed, and if the frequency characteristics are unsatisfactory, then a forced optimization of the coupler design is performed in order to obtain the required characteristics. The topology of the compact ring coupler is shown in Figure 5. The obtained frequency characteristics of the compact coupler in the AWR program are illustrated in Figures 6 and 7.

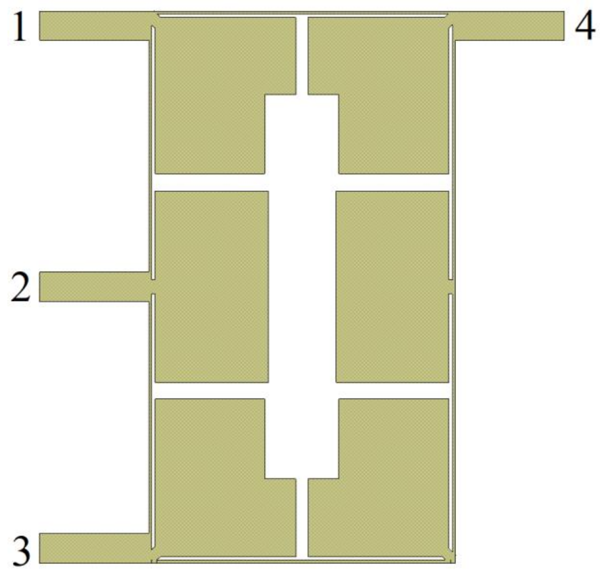


Fig. 5. Compact rat race coupler

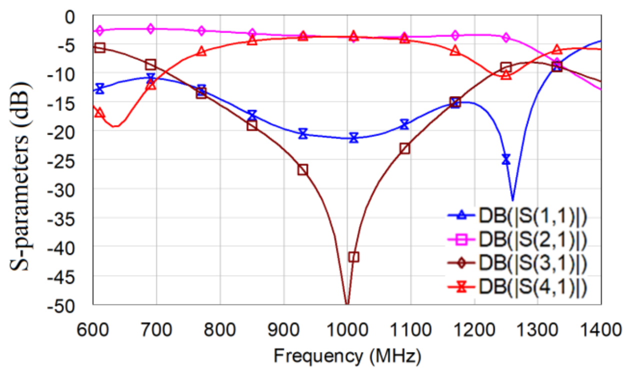


Fig. 6. S-parameter versus frequency plot for a compact rat race coupler

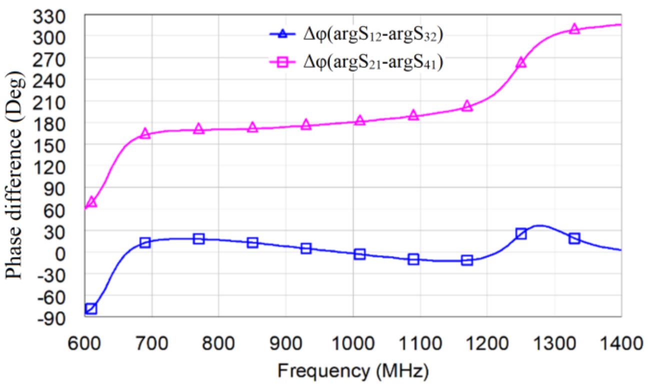


Fig. 7. The graph of the phase difference of the output signals depending on the choice of the input port (blue line - 1 input port, pink line - 4 input port)

By replacing the quarter-wave sections with low-pass filters and using the area inside the device, it was possible to reduce the area of the device by 82.3%. In a miniature version, the coupler has an area equal to 19.65 x 35.5 = 697.6 mm^2. According to the graphs obtained, it can be seen that the operating frequency band of the coupler at the isolation level "minus" 20 dB is 258 MHz (860-1118 MHz). The imbalance between the output signals of the structure does not exceed 1.1 dB. As a result of the miniaturization of the coupler, the operating frequency band was narrowed by 19.3%. The differences are primarily due to the incomplete coincidence of the characteristics of the quarter-wave sections and the elements installed instead of them. From the data in Table 1, you can compare the characteristics of the standard and miniature coupler.

TABLE I
COMPARATIVE DATA OF COUPLER

Parameters	Standard	Compact
bandwidth, MHz	320	258
Area, mm^2	3934,5	697,6
Relative area, %	100	17.7
Central frequency, MHz	1000	1000
The phase outputs on the central frequency,°	0 180	1.8 180.3

To increase the bandwidth of operating frequencies, an additional circuit is added to the design of the coupler, the length of which is equal to the wavelength in the line. To increase the bandwidth of the quadrature directional couplers, the addition of additional stub lines (cascading) is used. In the case of ring (common-phase-antiphase) taps, cascading is also used to increase the bandwidth. The dual ring coupler topology is shown in Figure 8. A single wavelength loop is added to the coupler design, with a characteristic impedance of 100 ohms. To further increase the strip, a third and subsequent circuit can be added, but this will lead to an increase in dimensions and complicate the design of the device. Low-pass filters are also asymmetrically designed to occupy the internal area of the coupler and further reduce the size of the device. The obtained frequency characteristics of the compact coupler in the AWR program are illustrated in Figures 9 and 10.

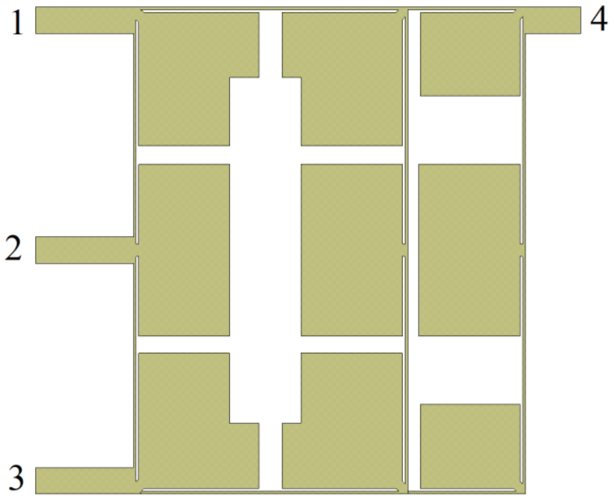


Fig. 8. Compact double rat race coupler

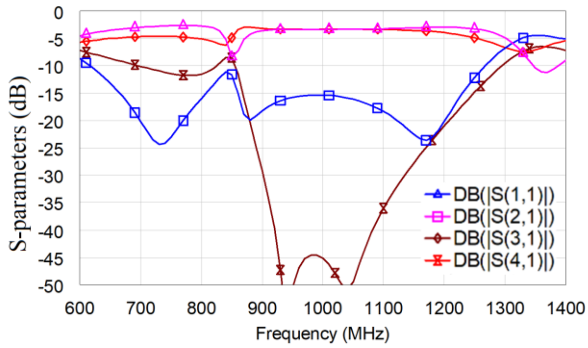


Fig. 9. S-parameter versus frequency plot for a compact double rat race coupler

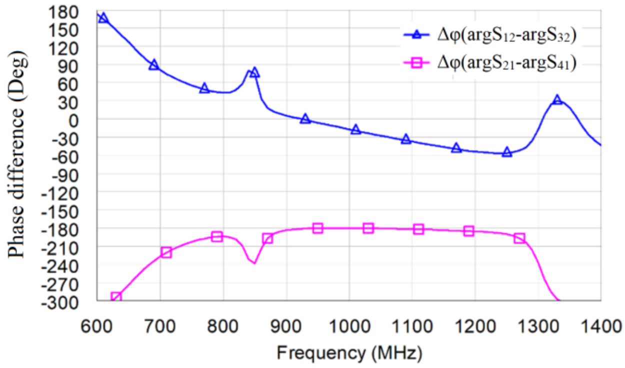


Fig. 10. The graph of the phase difference of the output signals depending on the choice of the input port (blue line - 1 input port, pink line - 4 input port)

By adding a loop, it was possible to increase the operating frequency band of the coupler to 329 MHz. At the same time, the area of the device increased to 977.2 mm^2. From the data in Table 2, you can compare the characteristics of the standard and miniature coupler.

TABLE II
COMPARATIVE DATA OF COUPLER

Parameters	Standard	Compact
bandwidth, MHz	465	329
Area, mm^2	6302	977.2
Relative area, %	100	15.5
Central frequency, MHz	1000	1000
The phase outputs on the central frequency,°	0 180	12 180

III. CONCLUSION

A miniature ring coupler obtained by replacing quarter-wave sections with low-pass filters is investigated in this work. Filter elements are located in the internal space of the device. All this made it possible to reduce the area of the coupler by 82.3% with a relative bandwidth of 25.8%, which is 19.3% less than that of a standard device. A double ring coupler was also obtained, the area of which is 84.5% less than the area of its standard implementation, with a relative bandwidth of 32.9%, which is 29.2% less than that of a standard device.

References

[1] Piotr Kurgan and Slawomir Koziel, "EM-driven compact cell topology selection for explicit size reduction of hybrid rat-race couplers," 2017 47th European Microwave Conference (EuMC). DOI: 10.23919/EuMC.2017.8230864.

[2] Vamsi Krishna Velidi, Sarika Shrivastava and Subrata Sanyal, "A compact-size microstrip rat-race coupler with high performance," IEEE Technology Students' Symposium. DOI: 10.1109/TECHSYM.2011.5783801.

[3] Po-Hsiang Tu and Chao-Hsiung Tseng, "Design of a compact rat-race coupler using dual transmission lines," 2010 Asia-Pacific Microwave Conference.

[4] Jianzhong Gu and Xiaowei Sun, "Miniaturization and harmonic suppression rat-race coupler using C-SCMRC resonators with distributive equivalent circuit," IEEE Microwave and Wireless Components Letters. DOI: 10.1109/LMWC.2005.859980.

[5] Slawomir Koziel and Adrian Bekasiewicz, "Novel structure and size-reduction-oriented design of microstrip compact rat-race coupler," 2016 IEEE/ACES International Conference on Wireless Information Technology and Systems (ICWITS) and Applied Computational Electromagnetics (ACES). DOI: 10.1109/ROPACES.2016.7465390.

[6] Tsu-Wei Lin, Jen-Yi Wu and Jen-Tsai Kuo, "Filtering rat-race coupler with transmission zeros using compact miniaturized hairpin resonators," 2015 IEEE International Wireless Symposium (IWS 2015). DOI: 10.1109/IEEE-IWS.2015.7164553.

[7] Zhebin Wang, Jae-Sik Jang and Chan-Wang Park, "Tri-band rat-race coupler using resonators," 2011 IEEE International Conference on Microwave Technology & Computational Electromagnetics. DOI: 10.1109/ICMTCE.2011.5915197.

[8] Saraunsh Bayaskar, Dewang Brahme and K Shambavi, "Compact rat race coupler with harmonic suppression using Circular defected ground structure (CDGS)," 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC). DOI: 10.1109/ICGTSPICC.2016.7955319.

[9] Chao-Hsiung Tseng and Hsiang-Ju Chen, "Compact Rat-Race Coupler Using Shunt-Stub-Based Artificial Transmission Lines," IEEE Microwave and Wireless Components Letters. DOI: 10.1109/LMWC.2008.2005225.

[10] Letavin Denis A., "Miniature microstrip rat-race couplers with artificial transmission lines," 2017 40th International Conference on Telecommunications and Signal Processing (TSP). DOI: 10.1109/TSP.2017.8076099.

[11] Tze-Min Shen, Chin-Ren Chen, Ting-Yi Huang and Ruey-Beei Wu, "Dual-band rat-race coupler design in multilayer LTCC," 2010 Asia-Pacific Microwave Conference.

[12] Chao-Hsiung Tseng, "Compact LTCC Rat-Race Couplers Using Multilayered Phase-Delay and Phase-Advance T-Equivalent Sections," IEEE Transactions on Advanced Packaging. DOI: 10.1109/TADVP.2010.2044660.

[13] Wei-Hsin Wang, Tze-Min Shen, Ting-Yi Huang and Ruey-Beei Wu, "Miniaturized rat-race coupler with bandpass response and good stopband rejection," 2009 IEEE MTT-S International Microwave Symposium Digest. DOI: 10.1109/MWSYM.2009.5165795.

[14] Gerard Siso, Jordi Bonache, Marta Gil, Joan Garcia-Garcia and Ferran Martin, "Compact Rat-Race Hybrid Coupler Implemented Through Artificial Left Handed and Right Handed Lines," 2007 IEEE/MTT-S International Microwave Symposium. DOI: 10.1109/MWSYM.2007.380209.

[15] Letavin D.A., “Bandwidth miniature quadrature branch line coupler design,” 2020 28th Telecommunications Forum, TELFOR 2020, November 2020. DOI: 10.1109/TELFOR51502.2020.9306606.

Experimental Analysis of Electromagnetic Interferences Absorber Influence on Metal Enclosure Immunity

Nataša J. Nešić, Slavko S. Rupčić, *Member, IEEE,* Vanja Mandrić-Radivojević, *Member, IEEE,* and Nebojša S. Dončov, *Senior Member, IEEE*

***Abstract*—This paper considers the experimentally conducted analysis of shielding effectiveness of an enclosure with electromagnetic interferences absorber placed at different positions inside. Dimensions of an absorber sheet are fitted to affect the enclosure's first resonant frequency in order to improve its protective function. An absorber sheet inside measured enclosure has an impact on the shielding effectiveness values not only at the first resonance but also in a wider frequency range.**

***Index Terms*—Absorber, Enclosure, EMI absorber sheet, Measurements, Shielding Effectiveness.**

I. Introduction

A common technique to protect electronic equipment from an external electromagnetic (EM) field and an electromagnetic interference (EMI) as well to limit a level of EM field, caused by equipment itself and radiated to surroundings, is to apply shielding [1]. A shielding enclosure may be constructed using steel, aluminum, copper or any other highly conductive material. Still, there are a number of different multipath coupling mechanisms, such as through apertures and cables, which can reduce a protective function of enclosure, usually expressed as the shielding effectiveness (SE). The SE can be determined as logarithmic ratio of electric field with and without shielding enclosure, in the same probe point. Also, the shielding characteristics of an enclosure can indicate negative values of the SE, especially at resonant frequencies of enclosure. Thus, the useful frequency range, in which EM shielding is provided, might be compromised.

To improve the shielding properties, several techniques can be applied. For instance, in [2] and [3], the SE of enclosure was increased by using absorbers or conductive foam. In [4] and [5], the authors proposed to suppress the first resonance in a metal enclosure by putting small antenna elements with loaded resistance. By placing a small dipole or loop antenna structure on the enclosure wall opposite to the enclosure aperture, it was shown that the EM shielding could be improved. In addition to that, it was numerically demonstrated in [6] and the experimentally confirmed in [7] that the physical dimensions of receiving-antenna, often used in experimental set-up to measure a level of EM field, could also affect the SE of enclosure.

In this paper, an influence of an EMI absorber on the SE of enclosure having the frontal wall with rectangular aperture is experimentally studied. During this study, this thin EMI absorber sheet is placed on one or more inner enclosure walls with the goal to consider how the position of absorber affects the SE behavior at the first resonance of enclosure but also at higher resonances. Some other EMI absorber positions inside the same considered enclosure are discussed in [9]. The paper is organized as follows. Section II refers to a physical enclosure's model with the EMI absorber material and a receiving-antenna inside. In Section III, the experimental set-up and measurement procedure are described. Section IV provides discussion of the experimental results. Finally, Section V summarizes the work.

II. Physical Model of Enclosure

The physical enclosure, which is used in the experimental measurements, with the monopole-receiving antenna and the absorber inside is shown in Fig.1. It has a front wall with rectangular aperture (not shown in Fig.1). The enclosure walls are made from copper material. The thickness of the enclosure walls is t = 1.5 mm. The metal enclosure entitled by D has a rectangular shape, with internal dimensions of (300x300x120) mm^3. On the front wall of the enclosure, a slot of dimensions (100x5) mm^2 is positioned symmetrically around the center on the frontal wall. An in-house monopole antenna is placed in the middle of the enclosure to measure the level of EM field inside. An antenna is made of copper wire with a length of 60 mm and radius of 0.1 mm. In order to prevent EM wave leakage from the tested enclosure, a copper strip is bonded to the joints.

Specifically, the absorber material is thin as a sheet of paper and it does not occupy significant space inside the enclosure. The 3M™ EMI Absorber AB7050 from AB7000 Series was available and therefore, it is used in this paper. One side consists of a flexible polymer resin loaded with soft metal flakes and on the other side is covered by an acrylic pressure-sensitive adhesive allows for easy application. This absorber is typically used for applications requiring EMI control and signal integrity improvement in the 50 MHz to 10

Nataša J. Nešić is with the Department of Information and Communication Technologies at Academy of Applied Technical and Preschool Studies Niš, Aleksandra Medvedeva 20, 18000 Nis, Serbia (e-mail: natasa.nesic@akademijanis.edu.rs).

Slavko M. Rupčić and Vanja Mandrić-Radivojević are with Department of Communications, Faculty of Electrical Engineering, Kneza Trpimira 2b, Osijek 31000, Croatia, (e-mails: rupcic@etfos.hr, vanja.mandric@etfos.hr).

Nebojša S. Dončov is with the Faculty of Electronic Engineering Niš, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia (e-mail: nebojsa.doncov@elfak.ni.ac.rs).

GHz range. It is a broadband EMI absorber designed to work in near-field applications inside and around electronic devices and assemblies [8]. Many factors determine the true attenuation of an EMI absorbing material, including shape and thickness, intimacy of substrate contact, smoothness of application surface, strength and frequency of the EMI signal, etc. [8].

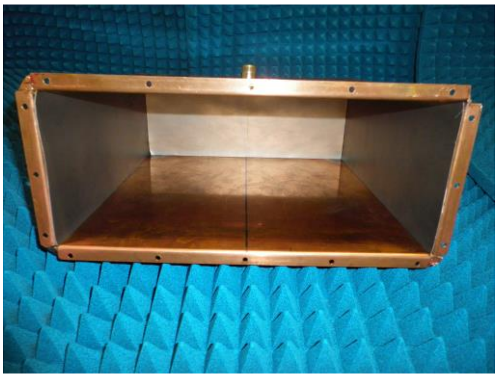


Fig. 1. Physical model of *D* metal enclosure with the EMI absorber on the wall opposite to the frontal enclosure wall and on both side walls.

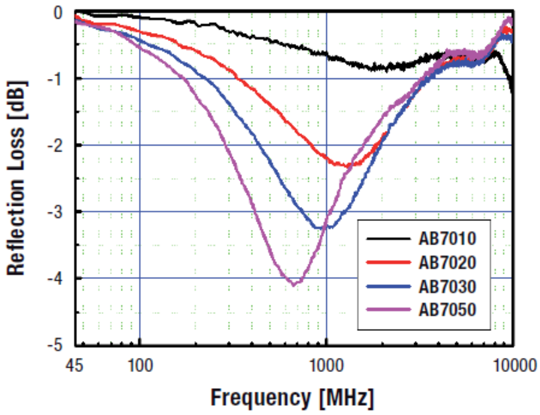


Fig. 2. The reflection loss of the EMI absorber AB7050 sheet which is used for measurement purposes, [8].

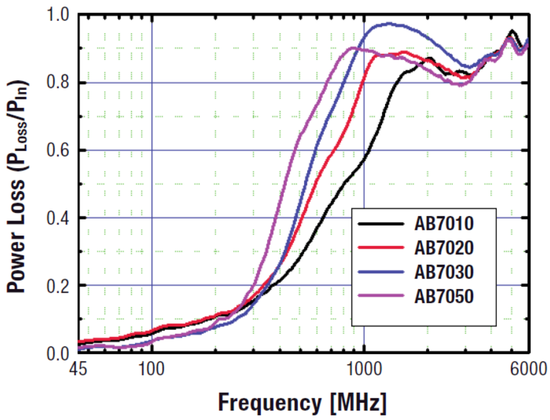


Fig. 3. The power loss of the EMI absorber AB7050 sheet which is used for measurement purposes, [8].

Figs. 2 and 3 present the reflection loss and the power loss of the EMI absorber, respectively, taken from [8]. The absorber characteristics and physical dimensions are given in Table I.

The EMI absorber sheet inside the enclosure *D* is placed in different positions in order to see how it will enhance the shielding characteristics of enclosure especially at the first resonant frequency. Firstly, the absorber is employed on the wall opposite to the front wall with an aperture which is entitled by BcW (back wall). Its dimensions correspond to the internal dimensions of the back enclosure's wall. Secondly, the absorbers are put on both side enclosure walls – entitled by 2SW (two side walls). The dimensions of the absorber are cut to fit the enclosure's side walls which is (297x120) mm^2. In the third case, the absorbers are employed at the same time on all above-mentioned positions, as depicted in Fig. 1. The SE characteristics obtained for this case will be called BcW+2SW.

TABLE I
THE EMI ABSORBER CHARACTERISTICS [8]

Property	Typical Value			
Adhesive	Acrylic non-conductive pressure-sensitive adhesive (PSA)			
Type of backing	Polymer resin with magnetic metal flake filter			
Product number	AB7010	AB7020	AB7030	**AB7050**
Backing thickness	0.1mm	0.2mm	0.3.mm	0.5mm
Adhesive thickness	0.03mm	0.05mm	0.05mm	0.05mm
Standard packaging	210 mm x 297 mm			
Temperature range	-25ºC ~ 85ºC			
Surface Resistivity	1 x 10^6 Ω (min)			
Initial permeability	110			
Typical performance	Refer to S_{11} attenuation and power loss graphs			

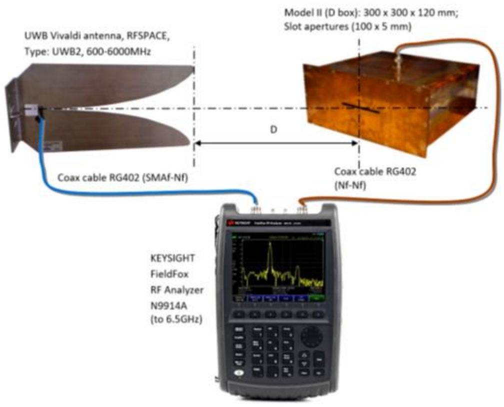


Fig. 4. The sketch of the measuring set-up: transmitting antenna, VNA and EUT (enclosure under test *D*).

III. EXPERIMENTAL PROCEDURE

The measurements are performed in a semi-anechoic room and the measuring set-up is illustrated in Figs. 4 and 5. The Keysight Field Fox RF Analyzer N9914A 6.5 GHz, with a resolution of 100 Hz and a maximum power of 3 dBm, is used as vector network analyzer (VNA). It is connected via cables to a transmitting antenna and a receiving antenna. A vertically polarized dipole antenna of type Vivaldi, as a broadband-antenna with a frequency range from 600 MHz to 6 GHz, is used as a transmitting antenna. As already mentioned, an in-house monopole antenna is used as a receiving antenna. All measurements are taken in the frequency range from 600 MHz to 2 GHz. The measurements are performed in the far-field. The distance is calculated to match that the enclosure under test (EUT) is in the far-field, which is marked as distance (*D*) in Fig. 4. The EMI absorber used in the experimental analysis is cut to fit the inner enclosure's sides.

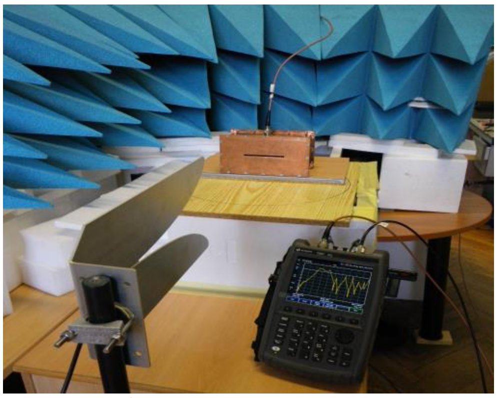


Fig. 5. The measuring configuration used in the semi-anechoic room.

The SE of enclosure is obtained by measure the electric field of EUT with monopole antenna inside and by measure the electric field of monopole antenna without enclosure, in the same probe point (the center of the considered enclosure) [7]. The SE of EUT is obtained by measuring the transmission parameters by using the network analyzer. It is measured twice, without and with enclosure which are marked as s_{21n} and s_{21e}, respectively. Therefore, the SE can be calculated by the following equation: SE [dB] = s_{21n} - s_{21e}.

IV. Discussion of Results

This section presents the experimental results of the SE of enclosure with the EMI absorber inside. To start with, the SE measurement results of enclosure *D* with the absorber employed on the wall opposite to the frontal wall with an aperture, are presented in Fig. 6. The SE results are obtained based on the measured level of EM field inside the enclosure. Figure 6 contains also the SE results for the case of enclosure without absorber (empty enclosure with only receiving antenna inside). As it can be seen from Fig. 6, a physical presence of the absorber led to the SE improvement of enclosure *D* in comparison to the empty enclosure case. In terms of shape, the compared SE curves are similar, however it can be observed that both curves differ regarding the SE levels not only around the first resonance but also at the second resonant frequency. The first resonance of empty enclosure is 686 MHz and the SE value is equal to -14.95 dB, which is critical and might compromise enclosure's shielding properties. On the other hand, the first resonance of enclosure with the EMI absorber occurs at 694 MHz and its SE has a positive value of 5.45 dB. The difference between the SE levels (ΔSE) at the first resonant frequency is 20.4 dB and the frequency shift (Δf_{r1}) related to the first resonance position without and with absorber is 8 MHz, i.e., the first resonant frequency is shifted toward higher frequencies when the EMI absorber is placed inside.

Secondly, the measurement results for the SE of enclosure with the EMI absorbers placed on both side walls are compared to the empty enclosure case. It can be seen that the SE curves do not differ in terms of shape in whole frequency range, see Fig. 7, but the SE levels are different at resonant frequencies around 700 MHz, 1100 MHz and 1650 MHz, respectively. It can be observed a very good absorber efficiency at lower frequencies, while it is weaker at higher frequencies in observed range. Since a very thin EMI absorber is used, the additional TE and/or TM modes are not introduced inside enclosure. The respective SE level difference between compared cases is 23.65 dB, while the frequency shift of the first resonance, Δfr_1, is 8 MHz.

Further, the EMI absorbers are placed at the same time on the wall opposite to the front wall with an aperture and on both side walls of enclosure. The measured SE results are compared to the empty enclosure case in Fig. 8. For this case, the highest SE values are obtained at resonant frequencies of enclosure. The difference between the SE levels (ΔSE) at the first resonance for this case and the case of empty enclosure is 26.5 dB, while the first resonant frequency shift, Δfr_1, is 10 MHz. Again, at higher frequencies, the insertion of absorbers for this case led to the decrease of the SE, as depicted in Fig. 8.

Finally, Fig. 9 presents the zoomed view of the SE around the first resonant frequency for all considered positions of EMI absorber. It can be observed that the SE enhancement is achieved for every position of absorber inside. Moreover, the more absorbers are placed inside the enclosure on different walls, the more significant improvement in the level of the SE characteristics is obtained, especially at the first resonant frequency.

Table II summarizes the measured values of the SE at the first resonant frequency as well as the relative shift of this resonance and SE difference from the case of enclosure without absorber. Definitely, the most prominent frequency shift is obtained for the third case, see Fig. 9. The highest SE value of 11.55 dB at the first resonant frequency is achieved for the case with three absorbers giving the difference of 26.5 dB compared to the SE of enclosure without absorber.

V. Conclusion

In this paper, the presence of the EMI absorber inside the metal enclosure has been considered from the viewpoint of the values of SE of enclosure at its resonant frequencies. The most significant improvement of SE value at the first resonant frequency is obtained for the case with absorbers on the back wall and two side enclosure walls. However, it should be taken into consideration that the EMI absorber presence may also reduce the SE peaks at the higher frequencies in observed frequency range.

Future work will include the numerical analysis together with a developed numerical model of used absorbers, to support the experimental study conducted here. Based on simulated distribution of the EM field inside the enclosure, it would be possible to determine the optimal position of the absorbers inside that can increase the protective function of enclosure at resonant frequencies but also keep its high SE values at other frequencies.

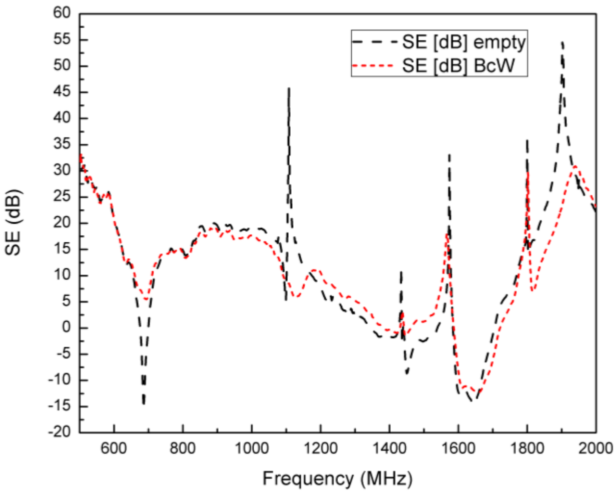


Fig. 6. The measurement results for the SE of the enclosure with the absorber placed on the wall on opposite side from frontal enclosure wall (case BcW).

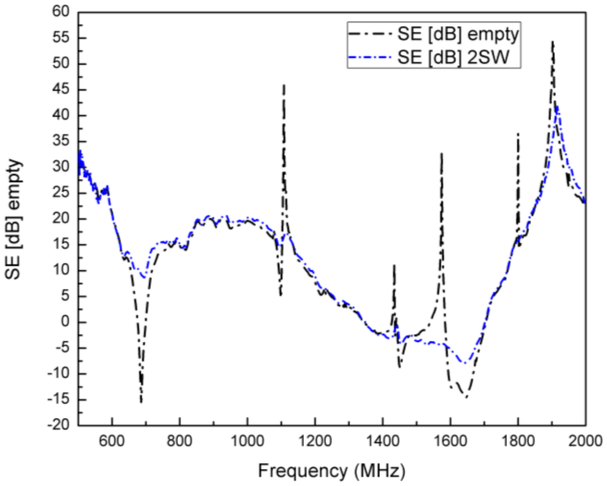


Fig. 7. The measurement results for the SE of the enclosure with the absorber placed on both side walls (case 2SW).

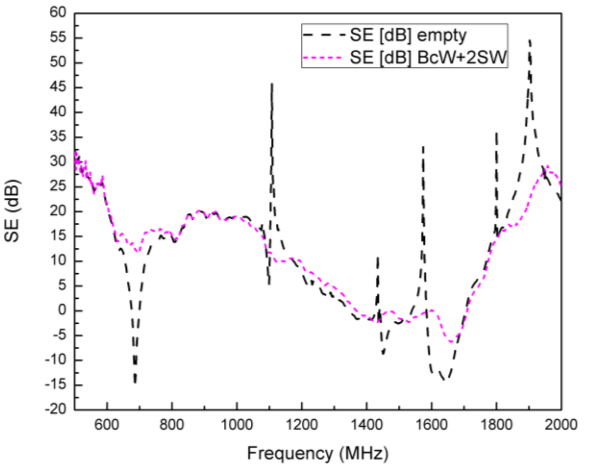


Fig. 8. The measurement results for the SE of the enclosure with the absorber placed on the wall on opposite side to the frontal wall and on both side walls (case BcW+2SW).

ACKNOWLEDGMENT

This work has been supported by the EUROWEB+ project, by the COST IC 1407 and by the Ministry of Education, Science and Technological Development of Republic of Serbia (Grant No. 451-03-9/2021-14/200102).

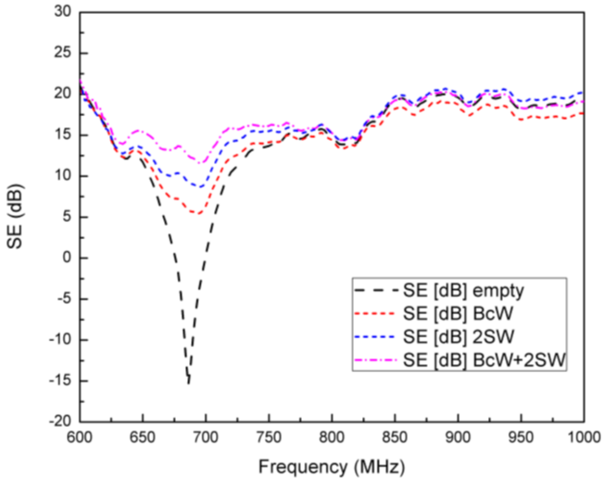


Fig. 9. The measurement results for the SE of enclosure, with the absorber inside, around the first resonant frequency (all three cases).

TABLE II
THE SE VALUES AT THE FIRST RESONANCE OF ENCLOSURE

EMI absorber position	fr_1_meas [MHz]	SE_meas [dB]	Δfr_meas [MHz]	ΔSE_meas [dB]
BcW	694	5.45	8	20.4
2SW	694	8.70	8	23.65
BcW+2SW	696	11.55	10	26.5
Empty	686	-14.95	-	-

REFERENCES

[1] C. Christopoulos, The Transmission-Line Modelling (TLM) Method, New Jersey: *IEEE Press in assoc. with Oxford University Press*, Piscataway, 1995.

[2] X. Luo and D. D. Chung, "Electromagnetic interference shielding using continuous carbon-fiber carbon-matrix and polymer-matrix composites," *Elsevier Science: Composites: Part B*, vol. 30, pp. 227–231, 1999.

[3] R. Kumar, S. R. Dhakate, P. Saini and R. B. Mathur, "Improved electromagnetic interference shielding effectiveness of light weight carbon foam by ferrocene accumulation," *The Roy. Soc. of Chem*. 2013: *RSC Advances*, vol. 3, pp. 4145-4153, 2013.

[4] N. Nešić, B. Milovanović, N. Dončov, V. Mandrić-Radivojević and S. Rupčić, "Improving Shielding Effectiveness of a Rectangular Metallic Enclosure with Aperture by Using Printed Dog-bone Dipole Structure," *52nd ICEST*, pp. 97-100, Niš, Serbia, June 28-30, 2017.

[5] Nataša Nešić, Bratislav Milovanović, Nebojša Dončov, Vanja Mandrić-Radivojević, Slavko Rupčić, "Improving Shielding Effectiveness of a Metallic Enclosure at Resonant Frequencies", *13th TELSIKS*, Niš, October 2017, pp. 42-45, ISBN: 978-1-5386-1798-4 (IEEE).

[6] N. J. Nešić and N. S. Dončov, "Shielding Effectiveness Estimation by using Monopole-receiving Antenna and Comparison with Dipole Antenna," *Frequenz DeGruyter*, no.5-6, pp. 1-11, 2016.

[7] N. J. Nešić "Numerical and experimental analysis of aperture arrays impact on the shielding effectiveness of metal enclosures in microwave frequency range", *Doctoral thesis, in Serbian, Singidunum University*, Belgrade, 2017.

[8] https://multimedia.3m.com/mws/media/960654O/3m-emi-absorber-ab7000hf-series-halogen-free.pdf

[9] N. Nešić, S. Rupčić, V. Mandrić-Radivojević, N. Dončov, "Experimental analysis of a metal enclosure shielding effectiveness improvement with EMI absorber", paper accepted for publication on *15th Conference on Applied Electromagnetics, ПЕС 2021*, Niš, Serbia.

Incorporating a Lowpass Filter into a Very Wide Bandpass Filter to Suppress Harmonics

Dušan Nešić, *Member, IEEE*

Abstract**—This paper presents an algorithm for suppression of higher harmonics in response of the very wide bandpass filter (WBPF). Lowpass filter (LPF) is incorporated into the very wide bandpass filter to suppress harmonics. The LPF is consists of only three identical cells with uniform open stubs. At least two higher harmonics are suppressed.**

Index Terms**— Microwaves, Suppression of higher harmonics for very wide bandpass filter, Ideal model.**

I. Introduction

ONE disadvantage of common bandpass filters is existing of higher harmonicas causing a narrow bandstop region on higher frequencies outside bandpass. This is specially a problem for the very wide bandpass filter (WBPF) like one presented in [1]. Common solution is incorporating lowpass filter (LPF) into the WBPF [2,3]. The solution intends not to significantly degrade bandpass characteristics but suppress higher harmonics. In the same time slow-wave characteristic of the lowpass filter shortens (minimizes) the structure. Problems can be nonuniform open stubs in [2] or too many different uniform open stubs like in [3].

Starting WBPF is from reference [1] with 150 % relative bandwidth. Ideal transmission structure and its response are presented in Fig. 1 and Fig. 2 in program package WIPL-D [4]. As can be seen in Fig. 2, bandstop region between the first bandpass and the next harmonic is narrow.

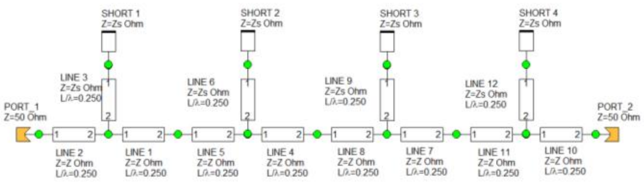


Fig. 1. Model of the ideal bandpass filter for relative bandwidth 150 %, presented in [1].

The aim is to suppress at least one higher harmonic. From Fig. 1 can be seen that transmission line segments between shorted stubs are λ/2 (2 x 0.250 λ), in phase π. To incorporate lowpass filter, the segment between the second and the third shorted stub is replaced with a lowpass filter. An example of a lowpass filter is chosen from [5]. It can be constructed of identical cells.

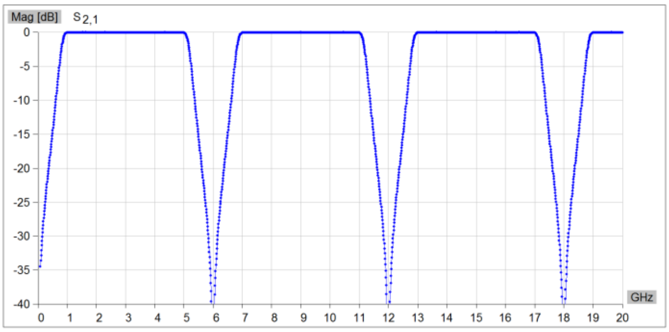


Fig. 2. S_{21} parameters for the filter around 3 GHz in the Fig. 1 with higher harmonica.

II. The Method

The method of incorporation of the lowpass filter is presented in Fig. 3. Instead of a transmission line segment π (λ/2 = 2 x 0.250 λ in upper Fig.3) between shorted stubs there are three equal cells of the lowpass filter, each with (*L1* + *L1*)* λ distance on the main line (3 x (*L1* + *L1*)* λ) in down Fig. 3). The phase difference π/3 is replaced with a cell presented in Fig. 4 with (*L1* + *L1*)* λ distance on the main line.

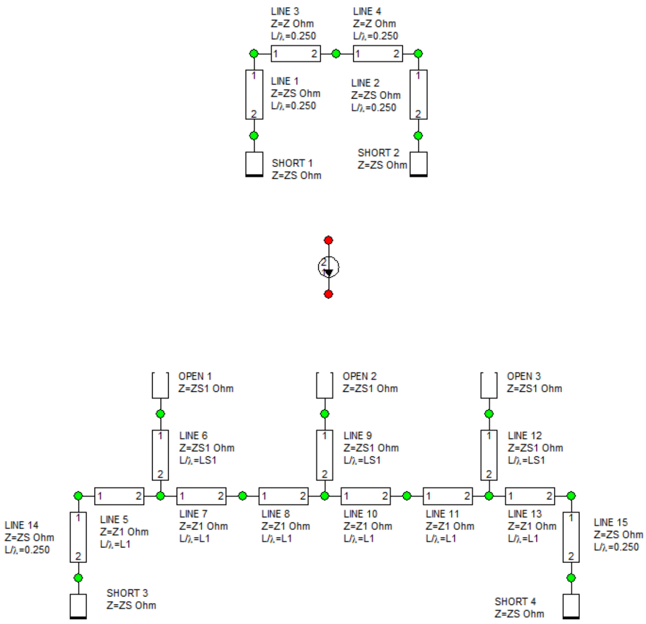


Fig. 3. Equivalence for incorporating low pass filter into bandpass filter.

Dušan Nešić is with Centre of Microelectronic Technologies, Institute of Chemistry, Technology and Metallurgy, University of Belgrade, Njegoševa 12, 11000 Belgrade, Serbia (E-mail: nesicad@nanosys.ihtm.bg.ac.rs).

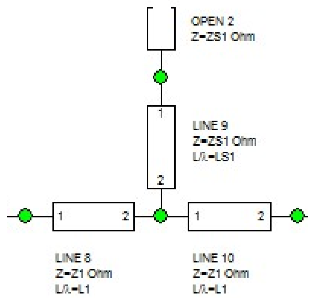


Fig. 4. One cell of the incorporated lowpass filter.

$A_cB_cC_cD_c$ matrix for the section π/3 of the primary filter with characteristic impedance *Z*, Fig. 1, is equaled on central frequency at 3 GHz with calculated ADCD matrix of the cell in Fig. 4. *L1* in Fig. 4 corresponds to phase difference *t1*, $L1 = t1/(2\pi)$. *Z1* is characteristic impedance of the main line and *ZS1* is characteristic impedance of the open stub. *LS1* corresponds to phase *tS1* of the open stub, $LS1 = tS1/(2\pi)$.

Calculation process is presented in Fig. 5 according to the reference [5], $Y = C_s/A_s$. Entering parameters are *Z*-characteristic impedance of the primary filter, *ZS*-characteristic impedance of the shorted stub of the primary filter (not in calculation, only in final results), *ZS1*-characteristic impedance of the open stub. *tS1*- phase length of the open stub was firstly treated as an entering parameter but later it is equaled with *t1* to get better response of the lowpass filter according to [5]. Output parameters are *t1* (*L1*) and *Z1*.

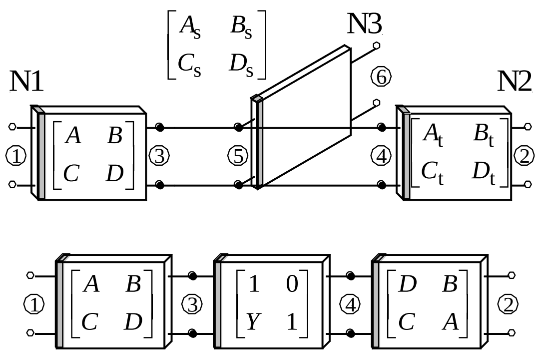


Fig. 5. Transformation of the cell in the Fig. 1 via ABCD matrix according to [4].

Equaled elements of matrix are presented in (1)-(3).

$$A_c(\pi/3) = \cos(\pi/3) = D_c(\pi/3) = AD + BC + BC_s \frac{D}{A_s} \quad (1)$$

$$B_c(\pi/3) = jZ\sin(\pi/3) = 2AB + C_s \frac{B^2}{A_s} \quad (2)$$

$$C_c(\pi/3) = j\sin(\pi/3)/Z = 2CD + C_s \frac{D^2}{A_s} \quad (3)$$

The corresponding ABCD-parameters for the cell in Fig. 4 are: $A = D = \cos(t1)$, $A_s = \cos(tS1)$, $B = \mathrm{j}Z1\sin(t1)$, $C = \mathrm{j}\sin(t1)/Z1$ and $C_s = \mathrm{j}\sin(tS1)/ZS1$.

Solution for $Z1$ from (1) and from (3) are (4) and (5) respectively

$$Z1 = \frac{ZS1\cdot(2\cos(2\cdot t1)-1)}{\sin(2\cdot t1)\cdot tg(tS1)} \quad (4)$$

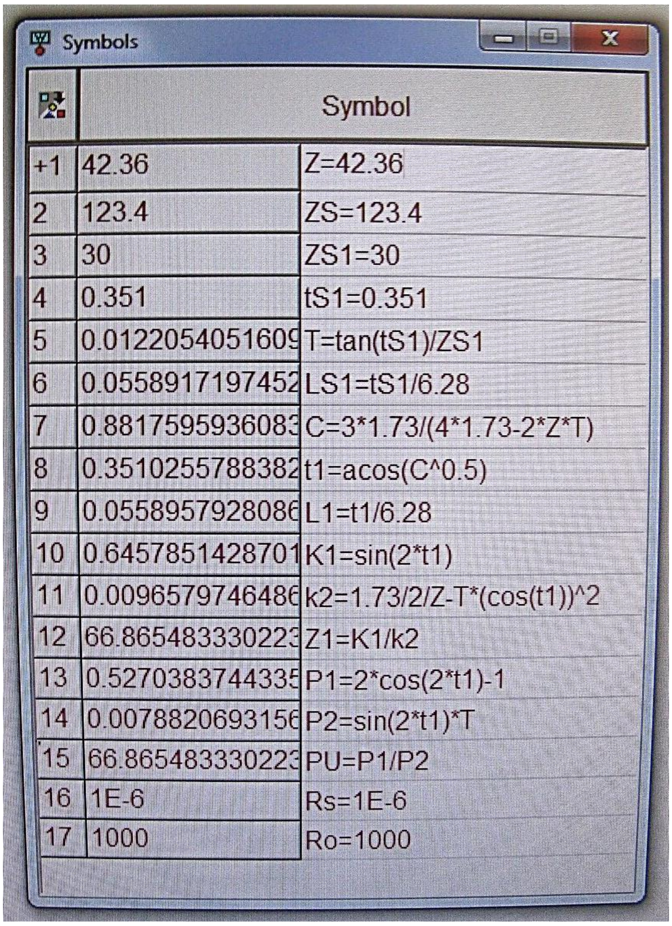
$$Z1 = \frac{\sin(2\cdot t1)}{(\sqrt{3}/(2\cdot Z) - (\cos(t1))^2 \cdot tg(tS1)/ZS1)} \quad (5)$$

Next, equal solutions from (4) and (5). After rearranging

$$(\cos(t1))^2 = 3\sqrt{3}/(4\sqrt{3} - 2Z(tg(tS1)/ZS1)) = C \quad (6)$$

$$t1 = a\cos(\sqrt{C}) \quad (7)$$

Return (5) to solution for $Z1$ and calculate $Z1$. Calculation in WIPL-D Microwave Pro v5.1 [4] is presented in Fig. 6. $Z1$ is firstly calculated via (6) and then checked via (5) (*PU*).



+1	42.36	Z=42.36
2	123.4	ZS=123.4
3	30	ZS1=30
4	0.351	tS1=0.351
5	0.0122054051609	T=tan(tS1)/ZS1
6	0.0558917197452	LS1=tS1/6.28
7	0.8817595936083	C=3*1.73/(4*1.73-2*Z*T)
8	0.3510255788382	t1=acos(C^0.5)
9	0.0558957928086	L1=t1/6.28
10	0.6457851428701	K1=sin(2*t1)
11	0.0096579746486	k2=1.73/2/Z-T*(cos(t1))^2
12	66.865483330223	Z1=K1/k2
13	0.5270383744335	P1=2*cos(2*t1)-1
14	0.0078820693156	P2=sin(2*t1)*T
15	66.865483330223	PU=P1/P2
16	1E-6	Rs=1E-6
17	1000	Ro=1000

Fig. 6. Calculation of parameters in WIPL-D Microwave Pro v5.1. *T*=tg(*tS1*)/*ZS1*. PU only checks calculation.

III. RESULTS

Entering values for *Z* and *ZS* are chosen according to [1] for the very wide bandpass filter with relative bandwidth of 150

% around 3 GHz. The value of characteristic impedance of the open stub, *ZS1*, is chosen to induce equality of *tS1* and *t1* for better characteristics of the lowpass filter [5].

The values from Fig. 6 are incorporated into models in Fig. 7. The final filter is symmetrical as can be seen in Fig. 7. Additional resistivity is incorporated to suppress numerical problems, Fig. 6. Results for S_{21} and S_{11} for the final structure with the incorporated lowpass filter are presented in Fig. 7. More than two higher harmonics are suppressed. S_{11} is below -15 dB in the bandpass.

IV. Conclusion

The algorithm for incorporating lowpass filter into bandpass is presented. For example, it is applied to a very wide bandpass (WBPF) filter of 150 %. Such WBPF has narrow bandstop and wide higher harmonic that make problems in circuits. More than two higher harmonics are suppressed with incorporated lowpass filter. The lowpass filter consists of only three identical cells with uniform arms in the middle of the WBPF structure. The algorithm is applicable to other bandpass filters. Further research will include simulation and fabrication in microstrip.

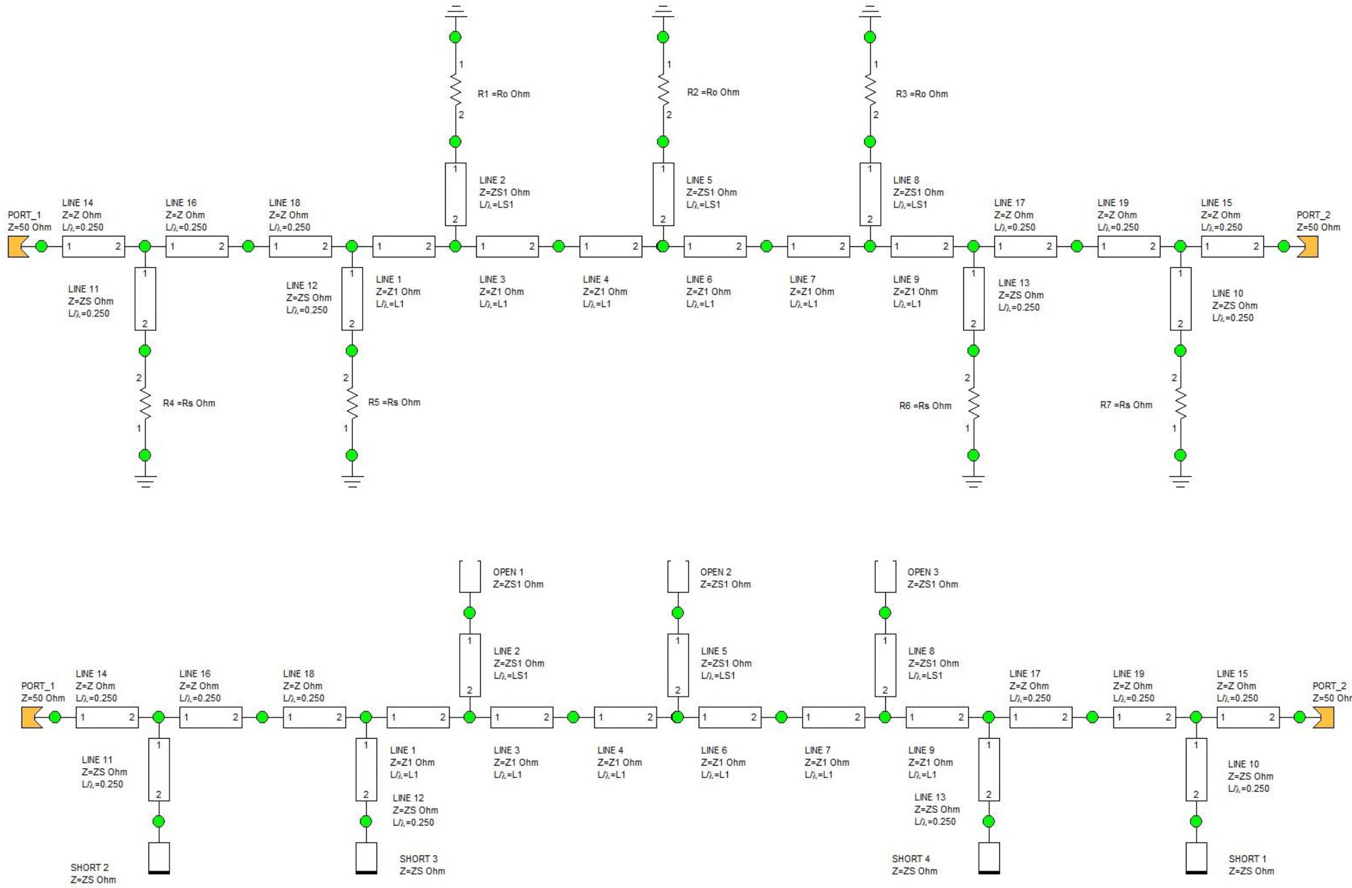


Fig. 7. Model of the lowpass filter incorporated into the bandpass filter, down model; the model with additional resistivity, upper model.

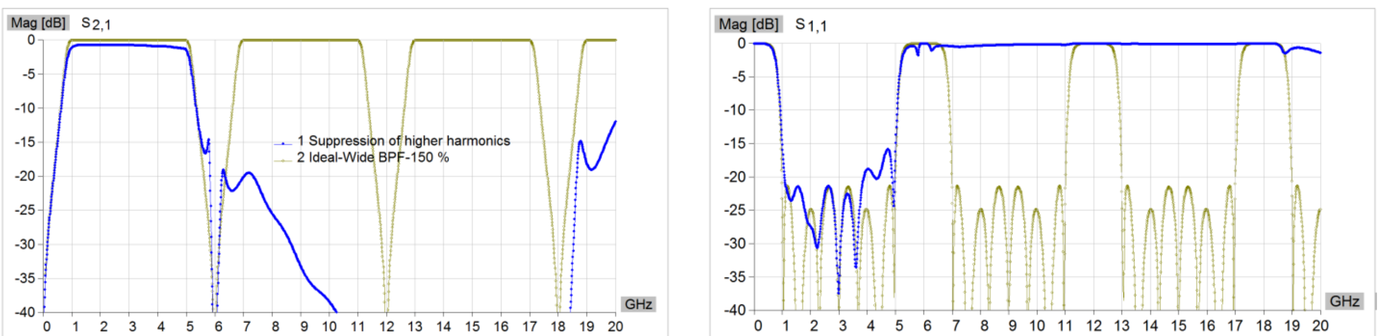


Fig. 8. S_{21} and S_{11} parameters for the resulting bandpass filter for *ZS1*-characteristic impedance of the open stub, equals 30 Ω.

ACKNOWLEDGMENT

This work was financially supported by Ministry of Education, Science and Technological Development of the Republic of Serbia (Grant No. 451-03-68/2020-14/200026).

REFERENCES

[1] D. Nesic, B. Kolundzija and T. Milosevic, “Simulation of wideband bandpass filters with arbitrary relative bandwidth,” *7th International Conference on Electrical, Electronic and Computing Engineering, IcETRAN 2020*, API1.8 (MTI)

[2] J. García-García, J. Bonache, and F. Martín, Application of Electromagnetic Bandgaps to the Design of Ultra-Wide Bandpass Filters with Good Out-of-Band Performance, *IEEE Transactions on Microwave Theory and Techniques*, Vol. 54, No. 12, 2006, DOI 10.1109/TMTT.2006.886155, pp. 4136-4140

[3] K.-K. Ryu, Bandpass Filter Using Short Stubs and Step Impedances Lines for Ultra-Wideband Applications, *Microwave and Optical Technology Letters*, Vol. 53, No. 5, 2011, DOI 10.1002/mop, pp. 1062-1065

[4] *Program Package* WIPL-D Microwave Pro v5.1 (*WIPL-D d.o.o, Belgrade 2019*. *www.wipl-d.com*)

[5] D. A. Nesic, B. M. Kolundzija, D. V. Tosic and D. S. Jeremic, Low-pass filter with deep and wide stop band and controllable rejection bandwidth, *International Journal of Microwave and Wireless Technologies*, Vol. 7, Iss. 2, 2015, pp. 141-149

НОВИ МАТЕРИЈАЛИ
/
NEW MATERIALS IN ELECTRICAL AND ELECTRONIC ENGINEERING

(НМ/NMI)

Uticaj jona retkih zemalja (Er, Yb, Ho) na karakteristike $BaTiO_3$ keramike

Vesna Paunović, *Member, IEEE*, Vojislav Mitić, Zoran Prijić, *Member, IEEE*

***Apstrakt* – U ovom radu, oksidi retkih zemalja, Er_2O_3, Yb_2O_3 i Ho_2O_3, su korišćeni kao materijali za dopiranje barijum-titanatne keramike. $BaTiO_3$ keramika dopirana je sa 0.01, 0.5 i 1.0 wt% Er, Yb ili Ho i pripremljena uobičajenim postupkom sinterovanja u čvrstom stanju. Uzorci su sinterovani na 1320 °C četiri sata. SEM analiza je pokazala da se u uzorcima $BaTiO_3$ dopiranim jonima retkih zemlje niskih koncentracija (0.01 wt%), veličina zrna kretala između 10 i 30 μm. Sa povećanjem koncentracije dopanata od 1.0 wt %, abnormalni rast zrna je zaustavljen, a veličina zrna se kretala između 2-10 μm. Merenja dielektrične konstante i dielektričnih gubitaka u zavisnosti od frekvencije i temperature vršena su u cilju uspostavljanja korelacije između mikrostrukture i dielektričnih svojstava dopirane $BaTiO_3$ keramike. Amfoterno ponašanje jona retkih zemalja dovodi do povećanja dielektrične konstante i smanjenja dielektričnih gubitaka u odnosu na nedopiranu $BaTiO_3$ keramiku. Ispitivana je i temperaturna zavisnost dielektrične konstante u zavisnosti od vrste i količine dopanata.**

Ključne reči – retke zemlje, $BaTiO_3$, dielektrična konstanta, Kirijeva temperatura, Kirijeva konstanta.

I. Uvod

Zbog visoke dielektrične konstante, temperaturne stabilnosti i malih gubitaka, elektrokeramika na bazi barijum titanata se široko koristi kao dielektrični materijal za višeslojne keramičke kondenzatore (MLCC), piezoelektrične senzore, aktuatore i DRAM memorije u integrisanim kolima [1-5]. S obzirom da su aditivi donorskog i akceptorskog tipa osnovne komponente dielektričnih materijala na bazi $BaTiO_3$, sprovedene su opsežne studije o njihovom uticaju na strukturu i svojstva $BaTiO_3$ [6-9]. Za dopiranje $BaTiO_3$ se mogu koristiti dve vrste dopanata: joni sa većim jonskim radijusima i valencom 3+ i većom koji mogu zameniti jone Ba^{2+} i joni sa manjim jonskim radijusima i valencom 5+ i većom koji mogu da zamene Ti^{4+} jone u perovskitnoj podrešetki [10-12]. U osnovi, uticaj jona dopanata na svojstva čvrstog rastvora zavisi od mesta koje dopantni jon zauzima u perovskitnoj strukturi, mehanizma kompenzacije i rastvorljivosti [13]. Pokazalo se da trovalentni joni supstituisani na Ba^{2+} lokacijama deluju kao donori, dok trovalentni joni supstituisani na Ti^{4+} lokacijama deluju kao akceptori. Kod ugradnje jona retkih zemalja u rešetku $BaTiO_3$, promena u strukturi i svojstvima $BaTiO_3$, uglavnom zavisi od mesta u rešetki gde je jon zamenjen [14].

Vesna Paunović – Univerzitet u Nišu, Elektronski fakultet, Aleksandra Medvedeva 14, 18000 Niš, Srbija (email: vesna.paunovic@elfak.ni.ac.rs).

Vojislav Mitić – Univerzitet u Nišu, Elektronski fakultet, Aleksandra Medvedeva 14, 18000 Niš, Srbija (email: vmitic.d2480@gmail.com).

Zoran Prijić – Univerzitet u Nišu, Elektronski fakultet, Aleksandra Medvedeva 14, 18000 Niš, Srbija (email: zoran.prijic@elfak.ni.ac.rs).

Joni iz sredine reda retkih zemalja, kao što su Yb^{3+}, Er^{3+}, Ho^{3+}, Dy^{3+}, $Sm^{3}+$, čiji su jonski radijusi po veličini između jonskih radijusa Ba^{2+} i Ti^{4+} jona, pokazuju amfoterno ponašanje i mogu zauzeti oba mesta katjonske rešetke u strukturi $BaTiO_3$ [13]. Iako neki teorijski proračuni sugerišu da će se, u širokom srednjem opsegu jonskih radijusa, trovalentni dopanti podeliti podjednako na dva mesta da bi se stvorila kompenzacija donor-akceptor [15], eksperimentalni rezultati pokazuju da se takva vrsta inkorporacije javlja u vrlo malo slučajeva [16]. Utvrđeno je da odnos Ba/Ti takođe utiče na ugradnju jona retkih zemalja u rešetku barijum titanata [17]. Ispitivanja uticaja retkih zemalja na strukturu $BaTiO_3$ pokazali su da se za odnos Ba/Ti> 1, joni Er^{3+}, Yb^{3+} ili Ho^{3+} ugrađuju na mesto Ti, a za odnos Ba/Ti <1 na mesto Ba. Štaviše, supstitucija jona retkih zemalja na mesta Ba ili Ti dovodi do neusklađenosti naelektrisanja sa rešetkom koje mora biti kompenzovano formiranjem negativno naelektrisanih defekata da bi se postigla ukupna elektroneutralnost [18]. Delimična zamena Ba^{2+} jona jonima retkih zemalja omogućava uniformnost mikrostrukture i sprečava abnormalni rast zrna.

Uzimajući u obzir da je kontrola mikrostrukture važna za optimizaciju električnih svojstava ovih materijala, u ovom radu ispitivan je uticaj Er_2O_3 i Yb_2O_3 i Ho_2O_3 na mikrostrukturu $BaTiO_3$ keramike i odgovarajuća električna svojstva.

II. Eksperimentalni deo

Uzorci su pripremljeni od komercijalnog praha $BaTiO_3$ (MURATA) i reagensa Er_2O_3, Yb_2O_3 i Ho_2O_3 u prahu (Fluka chemika). Koncentracija dopanata bila je 0.01, 0.5 i 1.0 wt% Er, Yb ili Ho. Uzorci su označeni kao 0.01Er-$BaTiO_3$ ($BaTiO_3$ dopiran sa 0.01 wt% Er) i tako dalje. Početni prah je mleven u mlinu sa kuglama u etil alkoholu tokom 24 sata koristeći boce od polipropilena i kugle od cirkonijuma. Posle nekoliko sati sušenja na 200°C, praškovi su presovani u diskove prečnika 7 mm i debljine 3 mm pod pritiskom od 120MPa. Kompakti su sinterovani na 1320°C u vazduhu tokom četiri sata. Za ispitivanje mikrostrukture uzorci su nagrizani u 10% HCl sa 5% HF. Mikrostrukture sinterovanih ili hemijski nagrizanih uzoraka posmatrane su skenirajućim elektronskim mikroskopom (JEOL-JSM 5300) opremljenim spektrometrom (EDS-QX 2000S). Raspodela veličine zrna i poroznosti uzoraka dobijena je pomoću LEICA K500MC sistema za obradu i analizu slika. Metoda linearnog merenja preseka korišćena je za procenu vrednosti veličine zrna, kao i odnosa zapremine pora. Pre električnih merenja, srebrna pasta je nanošena na ravne površine uzoraka. Vrednosti kapaciteta i dielektričnih gubitka izmerene su pomoću Agilent 4284 LCR metra u frekvencijskom opsegu od 100Hz-20kHz. Dielektrična konstanta izračunata je iz kapacitivnosti, debljine uzorka i površine elektroda. Varijacija dielektrične konstante sa temperaturom merena je u temperaturnom intervalu od 20°C do 180°C. Dielektrični parametri izračunati su prema Kiri Vajsovom i modifikovanom Kiri Vajsovom zakonu.

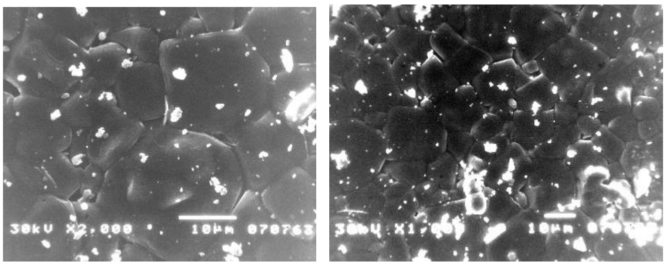
.

III. REZULTATI I DISKUSIJA

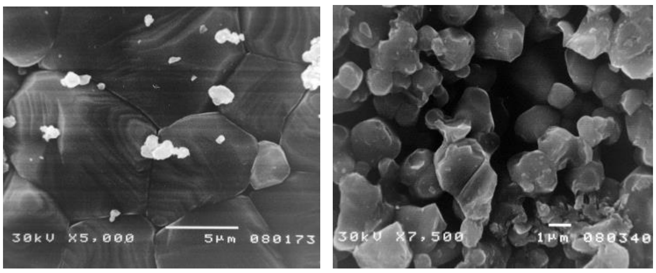
A. Mikrostrukturna svojstva

Ispitivanja gustine $BaTiO_3$ dopirane keramike su pokazala da za temperaturu sinterovanja od 1320°C gustina keramike varira od 82% teorijske gustine (TD), za uzorke dopirane većom koncentracijom aditiva, do 93% TD za uzorke dopirane nižom koncentracijom, pri čemu je najveća gustina izmerena za keramiku dopiranu Ho.

SEM analiza uzoraka dopiranih Er_2O_3 ili Yb_2O_3 pokazala je da ovi uzorci imaju sličnu mikrostrukturu. Uzorci dopirani Er_2O_3 poseduju zrna nepravilnog poligonalnog oblika (slika 1), dok su u Yb dopiranom $BaTiO_3$ karakteristična zrna sfernog oblika (slika 2). Spiralni koncentrični rast zrna primećen je u uzorcima dopiranim sa 0.01 wt% Er_2O_3 i Yb_2O_3. Ovaj oblik zrna se nije pojavljivao kod uzoraka sa većom koncentracijom aditiva. Za ove uzorke formiranje „staklaste faze“ ukazalo je da je sinterovanje izvršeno u tečnoj fazi. Za najnižu koncentraciju aditiva, veličina zrna se kretala do 30 μm. Sa povećanjem koncentracije dopanata veličina zrna se smanjivala.



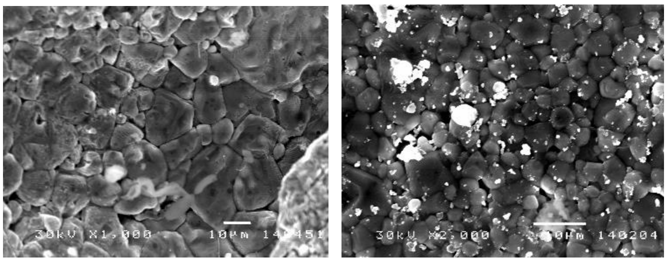
a) b)
Sl. 1 SEM mikrostruktura a) 0.01Er-$BaTiO_3$ i b) 1.0 Er-$BaTiO_3$.



a) b)

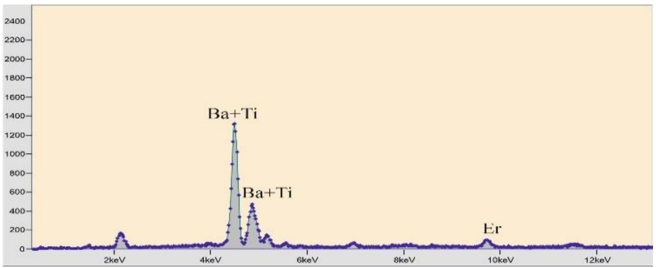
Sl. 2. Sl. 3. SEM mikrostruktura a) 0.01Yb-$BaTiO_3$ i b) 1.0 Yb-$BaTiO_3$.

Kao rezultat toga, za uzorke sa 0.5 wt% dopanta prosečna veličina zrna bila je od 10 do 15 μm, a za uzorke dopirane sa 1.0wt% aditiva veličina zrna bila je od 2-10μm.

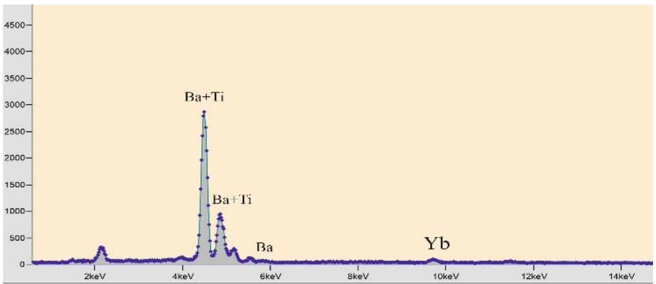


a) b)
Sl. 3. SEM mikrostruktura a) 0.01Ho-$BaTiO_3$ i b) 1.0 Ho-$BaTiO_3$.

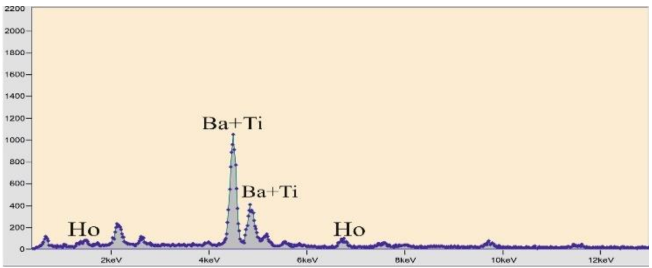
Uzorke keramike $BaTiO_3$ dopirane Ho_2O_3 karakterišu sferna i nepravilna poligonalna zrna (sl.3). Prosečna veličina zrna za uzorke dopirane niskim sadržajem Ho_2O_3 (0.01 wt% Ho) kretala se u rasponu od 10 do 20 μm (slika 3a). Povećavanjem koncentracije dopanata veličina zrna se smanjuje i za uzorke dopirane sa 0.5 wt% Ho veličina zrna se kretala od 2 μm do 10 μm, a za uzorke dopirane sa većom koncentracijom dopanata (1.0 wt% Ho) prosečna veličina zrna je bila od 2 μm do 3 μm (slika 3b).



a)



b)



c)

Sl. 4. EDS spektar a) Er-$BaTiO_3$, b) Yb-$BaTiO_3$ i c) Ho-$BaTiO_3$.

EDS analiza uzoraka dopiranih sa 0.01 wt.% dopanata nije pokazala postojanje oblasti bogatih Er, Yb ili Ho što je ukazivalo na uniformnu ugradnju dopanata u uzorke. Povećanje koncentracije dopanata dovelo je do pojave regiona bogatih Er između zrna kao i regiona bogatih Ho i Yb za koje je karakteristična sitnozrnasta mikrostruktura (slika 4).

Rentgenska analiza (XRD) uzoraka pokazala je samo $BaTiO_3$ perovskitnu fazu i ravnomernu raspodelu aditiva.

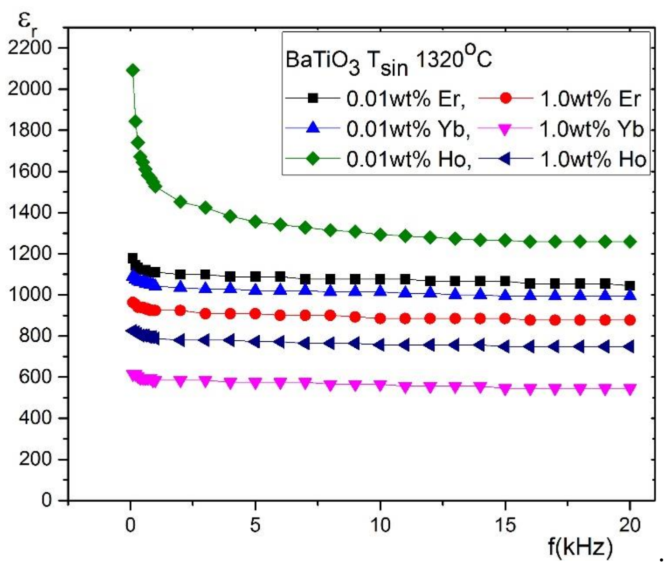
B. Električna svojstva

Uticaj koncentracije aditiva i dobijene mikrostrukture na dielektrične osobine uzoraka dopiranih jonima retkim zemalja se može ispitati zavisnošću dielektrične konstante i dielektričnih gubitaka od temperature i frekvencije (slike 5, 6 i 7). Frekventni opseg za sve ispitivane uzorke kretao se od 100Hz do 20kHz dok je temperaturni opseg bio od 20°C do 180°C.

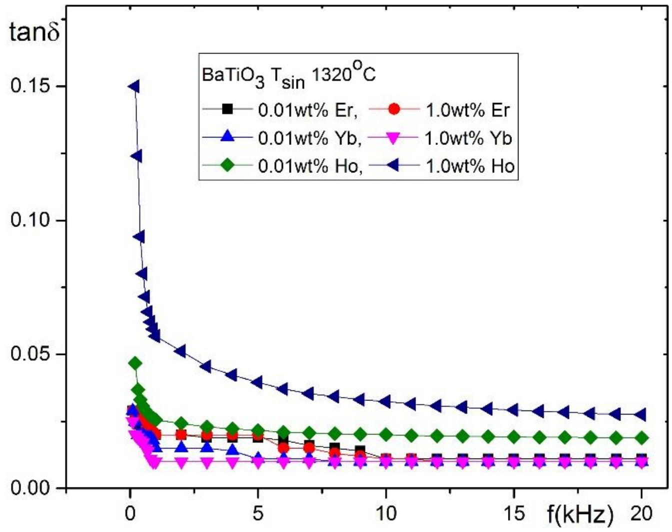
Tok promene dielektrične konstante sa frekvencijom je isti u svim uzorcima. Dielektrična konstanta nakon neznatno većih vrednosti na niskim frekvencijama opada i postaje gotovo konstantna za frekvencije veće od 3KHz.

Dielektrična konstanta ispitivanih uzoraka kretala se od 614 do 2100 na sobnoj temperaturi (Tabela I). Za 0.01-Er dopirani $BaTiO_3$ dielektrična konstanta iznosi 1200, a za 0.01-Yb-$BaTiO_3$, i 0.01-Ho-$BaTiO_3$ dielektrična konstanta iznosi 1150 i 2100 respektivno. Generalno, uzorci $BaTiO_3$ dopirani Ho pokazuju najveću dielektričnu konstantu u poređenju sa Er-$BaTiO_3$ i Yb –$BaTiO_3$ uzorcima.

Vrednosti dielektričnih gubitaka (*tan δ*) bile su u opsegu od 0.01-0.15. Glavna karakteristika svih uzoraka je da se nakon početno većih vrednosti dielektričnih gubitaka, *tan δ* smanjuje i gotovo je nezavisan od frekvencije pri frekvencijama većim od 5 kHz, što je prikazano na slici 6. Smanjenje *tan δ* u frekvencijskom opsegu od 100Hz-20 kHz može se povezati sa smanjenjem električne otpornosti uzoraka sa 10^6 Ωcm na 100 Hz na 10^2 Ωcm na 20 kHz.

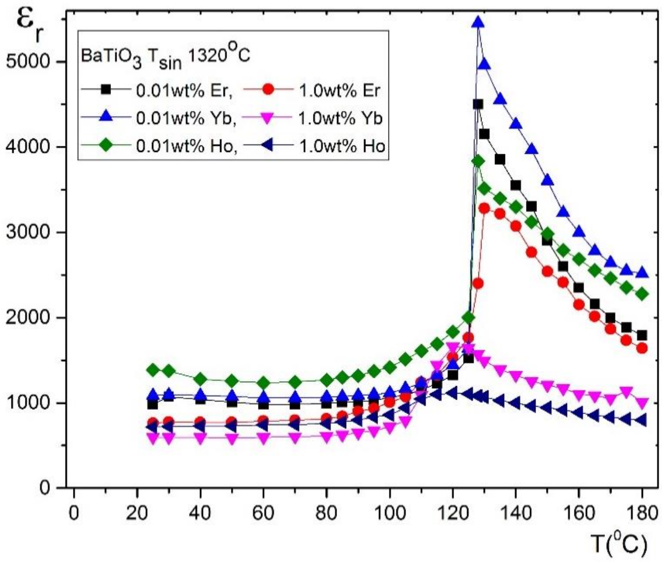


Sl. 5. Zavisnost dielektrične konstante od frekvencije.



Sl. 6. Zavisnost dielektričnih gubitaka od frekvencije.

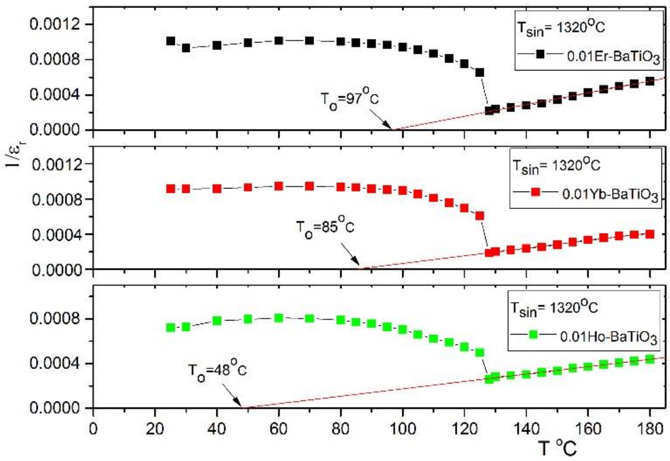
Uticaj tipa aditiva i karakteristika mikrostrukture na dielektrično ponašanje Er, Yb i Ho dopiranog $BaTiO_3$ može se proceniti pomoću krivih zavisnost permitivnosti od temperature (slika 7). Najveća promena dielektrične konstante u odnosu na temperaturu za uzorke sa niskom koncentracijom aditiva (0.01 wt%) primećena je u Yb dopiranom $BaTiO_3$ za koji je dielektrična konstanta na Kirijevoj temperaturi 5453. Relativno stabilan odziv dielektrične konstante u funkciji temperature do 100°C primećen je u svim dopiranim uzorcima. Sa većom koncentracijom aditiva (1.0 wt%), primećena je mala promena dielektrične konstante sa temperaturom u celom temperaturnom intervalu kod svih uzoraka. Smanjenje dielektrične konstante u dopiranim uzorcima sa porastom koncentracije dopanata, može se pripisati s jedne strane segregaciji dopanata, a sa druge strane smanjenju gustine uzoraka sa 93 na 82% TG. Kirijeva temperatura (T_C) na kojoj je vrednost ε_r maksimalna, niža je od vrednosti T_C za čisti $BaTiO_3$ i kretala se u opsegu od 122°C do 129°C.



Sl. 7. Zavisnost dielektrične konstante od temperature.

Dielektrična konstanta u feroelektriku raste sa temperaturom, dostiže maksimalnu vrednost na Kirijevoj temperaturi i opada sa daljim porastom temperature. Zavisnost dielektrične permitivnosti od temperature u paraelektričnom regionu, tj. u području iznad Kirijeve temperature, može se opisati Kiri-Vajsovim zakonom.

Fitovanjem zavisnosti recipročne vrednosti dielektrične konstante od temperature, kao što je prikazano na slici 8, dobijaju se vrednosti Kiri-Vajsove temperature T_0. Ova temperatura ima nižu vrednost od Kirijeve temperature T_C za sve izmerene uzorke. Najviša vrednost za T_0 dobijena je za uzorke dopirane sa 0.01 wt% Er (T_0 = 97°C), a najniža vrednost T_0 za uzorke dopirana sa 1.0 wt% Ho (T_0 = 10°C). U tabeli I date su vrednosti Kiri-Vajsove temperature za sve ispitivane uzorke.



Sl. 8. Recipročna vrednost dielektrične konstante u funkciji temperature.

Na osnovu Kiri-Vajsovog zakona izračunavaju se vrednosti Kirijeve konstante *C* za sve izmerene uzorke. Vrednost Kirijeve konstante opada sa porastom koncentracije aditiva tako da se najveća vrednost izračunava za uzorke sa koncentracijom aditiva 0.01 wt% ($3.04 \cdot 10^5$ K za Ho-$BaTiO_3$, $2.34 \cdot 10^5$K za Yb-$BaTiO_3$ i $1.53 \cdot 10^5$ K za Er-$BaTiO_3$), a najmanja za uzorke sa koncentracijom aditiva 1.0 wt%. Uzorke sa najvećom vrednošću *C* karakterišu krupnozrna mikrostruktura i veća gustina. Vrednosti za Kirijevu konstantu u skladu su sa promenom gustine ispitivanih uzoraka, kao i sa mikrostrukturnim karakteristikama.

Za sve izmerene uzorke karakterističan je oštar prelaz iz feroelektrične u paraelektričnu oblast, kao i nagli porast dielektrične permitivnosti na Kirijevoj temperaturi. Ova činjenica se može potvrditi odnosom dielektrične konstante na Kirijevoj temperaturi (ε_{rmax}) i na sobnoj temperaturi (ε_{rmin}), tj. (ε_{rmax} / ε_{rmin}). Najviši odnos dielektrične konstante izračunat je za uzorke dopirane sa 0.01 wt% (od 4.74 za 0.01-Yb-BaTiO do 1.61 za 0.01-Ho-$BaTiO_3$), a najmanji odnos uzoraka dopirane sa 1.0wt%.

Pored Kiri-Vajsovog zakona za ispitivanje ponašanja feroelektrika u paraelektričnoj fazi koriste se i modifikovani Kiri-Vajsov zakon koji opisuju odstupanja od linearnosti $\varepsilon_r = f(T)$ usled difuzne fazne transformacije:

$$\frac{1}{\varepsilon_r} = \frac{1}{\varepsilon_{r\max}} + \frac{(T - T_{\max})^{\gamma}}{C'}. \tag{1}$$

gde je C' konstanta slična Kirijevoj konstanti, γ - kritični eksponent nelinearnosti koji pokazuje odstupanje od linearne zavisnosti ε_r od temperature u paraelektričnoj oblasti. Na osnovu fitovanja krivih $ln(1/\varepsilon_r - 1/\varepsilon_{rmax})$ u funkciji $ln(T-T_{max})$, γ je dobijen kao nagib krive.

Vrednost kritičnog eksponenta nelinearnosti γ za izmerene uzorke kreće se od 1.02 za uzorke dopirane 0.01wt% Ho do 1.19 za uzorke dopirane 1.0wt% aditiva (Tabela I). Vrednost γ raste sa povećanjem koncentracije aditiva. Ove vrednosti su u skladu sa eksperimentalnim podacima, jer je za ove uzorke karakterističan oštar prelaz iz feroelektričnog u paraelektrični region što ukazuje na strukturnu faznu transformaciju.

TABELA I
DIELEKTRIČNI PARAMETRI ZA ISPITIVANE UZORKE

Uzorci $BaTiO_3$	ε_r na T=300K	ε_r na T_C	T_c [°C]	T_0 [°C]	$C \cdot 10^5$ [K]	γ
0.01wt %Er	1200	4505	128	97	1.53	1.12
0.5wt%Er	1199	4100	128	87	1.24	1.18
1.0wt%Er	1010	3281	129	85	1.65	1.19
0.01wt%Yb	1150	5453	128	85	2.34	1.11
0.5wt%Yb	614	1660	125	28	1.55	1.15
1.0wt%Yb	610	1657	125	27	1.48	1.19
0.01wt%Ho	2100	3834	128	48	3.04	1.02
0.5wt%Ho	800	1290	125	25	1.62	1.14
1.0wt%Ho	700	1117	122	10	1.54	1.17

IV. ZAKLJUČAK

U ovom radu su predstavljena ispitivanja uticaja jona retkih zemalja, kao što su Yb_2O_3, Er_2O_3 i Ho_2O_3 na mikrostrukturu $BaTiO_3$ keramike i odgovarajuća električna svojstva. Naša ispitivanja su pokazala da za temperaturu sinterovanja od 1320°C gustina keramike varirala od 82% teorijske gustine (TD), za uzorke dopirane većom koncentracijom dopanata do 93% TD za uzorke sa niskom koncentracijom dopanata. Primećeno je da povećanje sadržaja katjona retkih zemljanih sprečava abnormalni rast zrna. Prosečna veličina zrna u uzorcima dopiranim sa malim sadržajem aditiva kretala se između 10-30 μm, a kod 1.0 wt% od 2-10 μm. Dielektrična merenja su pokazala da, generalno, uzorci Ho-$BaTiO_3$ pokazuju najveću vrednost dielektrične konstante na sobnoj temperaturi a da su najveće promene dielektrične konstante sa temperaturom zabeležene kod Yb-$BaTiO_3$ uzoraka. Dielektrična konstanta ispitivanih uzoraka kretala se od 610 do 2100 na sobnoj temperaturi. Smanjenje dielektrične konstante sa povećanjem koncentracije dopanata povezano je sa neujednačenom raspodelom dopanata u strukturi perovskita. Svi uzorci imaju oštar fazni prelaz i slede Kiri-Vajsov zakon. Kirijeva konstanta (*C*) opada sa povećanjem količine aditiva za sve vrste uzoraka. Kritični eksponent γ je u rasponu od 1.02 do 1.19 i povećava se sa porastom koncentracije aditiva. Dobijeni rezultati omogućavaju dalju optimizaciju električnih svojstava materijala na bazi barijum-titanata, posebno sa stanovišta sinteze-struktura-svojstva.

ZAHVALNICA

Ovaj rad je podržalo Ministarstvo prosvete, nauke i tehnološkog razvoja Republike Srbije (Ev. br. 451-03-9/2021-14/200102).

LITERATURA

[1] S.Wang, G.O.Dayton "Dielectric Properties of Fine-Grained Barium Titanate Based X7R Materials" *J. Am. Ceram. Soc.* 82 [10], pp 2677–2682, 1999.

[2] C.Pithan, D.Hennings, R. Waser "Progress in the Synthesis of Nanocrystalline $BaTiO_3$ Powders for MLCC" *International Journal of Applied Ceramic Technology* 2 [1], pp.1–14, 2005.

[3] Y. Lu, H. Haoa, S. Zhang, H. Liu, C. Su, Z. Yao, M. Cao, J " Microstructure and dielectric characteristics of Nb_2O_5 doped $BaTiO_3$-$Bi(Zn_{l/2}Til_{/2})O_3$ ceramics for capacitor applications", . E. Ceram. Soc., 37, pp. 123-128, 2017.

[4] J.Nowotny, M.Rekas "Positive temperature coefficient of resistance for $BaTiO_3$-based materials" *Ceram. Int*., 17 [4], pp. 227-41, 1991.

[5] C.Metzmacher, K.Albertsen, "Microstructural Investigations of Barium Titanate-Based Material for Base Metal Electrode Ceramic Multilayer Capacitor" J.Am.Ceram.Soc. 84 [4], pp. 821-26, 2004.

[6] D.H. Kuo, C.H.Wang, W.P.Tsai, ″Donor and acceptor cosupstitued $BaTiO_3$ for nonreducible multilayer ceramic capacitors″, Ceram. Int. vol. 32, pp. 1-5, 2006.

[7] V. Paunovic, V. V. Mitic, Lj. Kocic, ″Dielectric characteristics of donor-acceptor modified $BaTiO_3$ ceramics″, Ceram. Int. vol. 42, no. 10, pp. 11692-11699, 2016.

[8] N.H.Chan, D.M.Smyth, "Defect chemistry of donor-doped $BaTiO_3$" *J. Am. Ceram. Soc*., 67 [4] pp. 285-8 1984.

[9] D.ltzke, H.Abicht, "The influence of different additives and the mode of their addition on the sintering behavior and the properties of semiconducting barium titanate ceramics" *Solid State Sciences* 2, pp.149–159, 2000.

[10] V. Paunovic, V. V. Mitic, M. Miljkovic, V. Pavlovic, Lj. Živkovic, ″Ho_2O_3 Additive Effects on $BaTiO_3$ Ceramics Microstructure and Dielectric Properties″ *Science of. Sintering*, vol. 44, no. 2, pp. 223-233, 2012.

[11] H.M.Chan, M.P.Hamer, D.M.Smyth, "Compensating defects in highly donor-doped $BaTiO_3$" *J.Am. Ceram. Soc*., 69 [6] 507-10, 1986.

[12] P.W.Rehrig, S.Park, S.Trolier-McKinstry, G.L.Messing, B.Jones, T.Shrout "Piezoelectric properties of zirconium-doped barium titanate single crystals grown by templated grain growth" *J. Appl. Phys.* 86 [3], 1657-1661, 1999.

[13] D.Makovec, Z.Samardzija M.Drofenik "Solid Solubility of Holmium, Ytrium and Dysprosium in $BaTiO_3$" *J.Am.Ceram.Soc.*, 87 [7], pp. 1324-1329 (2004).

[14] D.Lu, X.Sun, M.Toda, "Electron Spin Resonance Investigations and Compensation Mechanism of Europium-Doped Barium Titanate Ceramics" *Japanese Journal of Applied Physics* 45 [11] pp.8782-8788 (2006).

[15] M. T. Buscaglia, V. Buscaglia, P. Ghigna, M. Viviani, G. Spinolo, A. Testino and P. Nanni, "Amphoteric behaviour of Er^{3+} dopants in $BaTiO_3$: an Er–L_{III} edge EXAFS assessment" *Phys. Chem. Chem. Phys.*, 6**,** pp. 3710, 2004.

[16] G.V. Lewis, C.R.A. Catlow, " Defect studies of doped and undoped barium titanate using computer simulation techniques" *J. Phys. Chem. Solids* 47 pp. 89-97, 1986.

[17] L.Xue, Y.Chen, R.Brook, "The influence of ionic radii on the incorporation of trivalent dopants into $BaTiO_3$" *Materials Sci. and Eng. B* Vol. 1 2, pp. 193-201, 1988.

[18] John D. Bak, John C. Wright "Site-Selective Spectroscopy of the Solid-State Defect Chemistry in Erbium-Doped Barium Titanate" *J. Phys. Chem. B, 109,* 18391-18399, 2005.

ABSTRACT

In this paper, rare earth oxides, Er_2O_3, Yb_2O_3, and Ho_2O_3, were used as doping materials for barium titanate ceramics. $BaTiO_3$-ceramics were prepared by a conventional solid-state reaction method. The concentration of additives was 0.01, 0.5, and 1.0 wt%. The samples were sintered at 1320°C for four hours. In $BaTiO_3$ samples doped with rare-earth ions of low concentrations (0.01 wt%), the grain size ranged between 10 and 30 μm. With an increase in dopant concentration of 1.0 wt%, abnormal grain growth was inhibited, and grain size ranged between 2-10μm. Measurements of dielectric constant and dielectric losses as a function of frequency and temperature were performed to establish a correlation between the microstructure and dielectric properties of doped $BaTiO_3$ ceramics. The amphoteric behavior of rare-earth ions leads to an increase in the dielectric constant and a decrease in dielectric losses in relation to undoped $BaTiO_3$ ceramics. The temperature dependence of the dielectric constant in function on the type and concentration of dopants was also investigated.

Influence of rare earth ions (Er, Yb, Ho) on the $BaTiO_3$ ceramics characteristics

Vesna Paunović, Vojislav Mitić, Zoran Prijić

Mössbauer Spectroscopy of Iron-based Chalcogenides

Valentin N. Ivanovski

***Abstract*—Soon after the discovery of superconductivity in LaFeAsO with T_C = 26 K in 2008, many other iron-based superconductors were synthetized. They are all based on the layers which contain iron and a pnictogen (As, P) or a chalcogen (S, Se, Te) element. Due to the connection between superconductivity and magnetism these novel unconventional high-T_C superconductors have attracted tremendous interest in the scientific community. A particularly well studied is tetragonal FeSe in the PbO type structure (11 family). The improvement of T_C was achieved by the intercalation of an additional layer such as perovskite-like blocks or alkaline metals into the Fe-based chalcogenide layered systems. This led to creation of new superconducting compounds, $A_yFe_{2-x}Se_2$ (A is an alkaline element) named 122 family whose physical and structural properties are found to be very sensitive on details of the chemical composition. Unlike layered cuprate superconductors, a cationic disorder arisen from a substitution in an Fe-layer improves T_C. The highest T_C in the Fe-based chalcogenide superconductors is accomplished by suppression of both long range crystallographic and magnetic order. Mössbauer spectroscopy is a very useful tool for studies of structural phase transitions, structure defects, and chemical and structural inhomogeneities. This lecture is devoted to the local structure studies of $FeSe_xS_{1-x}$, $K_{0.7}Na_{0.1}Fe_2Se_2$, $KFe_{1-x}Co_xSe_2$, and similar Fe-based chalcogenide compounds using the Mössbauer spectroscopy.**

***Index Terms*—Superconductors; iron-based chalcogenides; Mössbauer spectroscopy; local structures.**

Valentin N. Ivanovski is with the “Vinča” Institute of Nuclear Sciences - National Institute of the Republic of Serbia, University of Belgrade, P.O.Box 522, 11001 Belgrade, Serbia (e-mail: valiva@vin.bg.ac.rs).

An Overview on a Graph Theory Applications New Frontiers in Electronics Materials

Vojislav V.Mitic, Branislav Randjelovic, Dusan Milosevic, Srdjan Ribar, Ivana Radovic, Markus Mohr and Hans Jorg Fecht

***Abstract* — This paper are new exciting results and an overview in graph theory applications on various problems in electronic materials. This opens new frontiers in this field. There are a lot of scientific efforts, especially during last few years. Our research team made very significant contribution, particulary related to calculation of parameters of BaTiO3 ceramics, calculation of syntetized diamonds electrophysical parameters, modelling of microeletronic intergranular relations and use of materials in medicine, based on biomimetic. We also discuss interesting possibilities for further research.**

***Index Terms* — graph theory; nanostructures; ceramics.**

I. Introduction

Graph theory application in synthesis of ceramic materials with control on phenomena between the grains is very challenging (see [1-5]), especially combined or compared with fractals and neural networks applications. It could enable design of the material with specific functionalities and widen the field and conditions for their application. Properties on the grain boundary can be calculated and analyzed based on the values measured on bulk materials samples (for methods of measurement see [1,3,4,5,8]). Possibility of using graph theory for determination of dielectric properties at the grain boundary of modified $BaTiO_3$ ceramics, based on the properties measured on the bulk samples is shown. This opens completely new perspectives in the area of miniaturization and micropackaging, because electrical-electronic-dielectric parameters values are calculated on a micro level (see [3, 6-9]).

Vojislav V.Mitić is with Faculty of Electrical Engineering, University of Nis, 14 Aleksandra Medvedeva, 18000 Belgrade, Serbia (e-mail: vmitic.d2480@gmail.com).

Branislav M.Randjelović is with Faculty of Electrical Engineering, University of Nis, 14 Aleksandra Medvedeva, 18000 Belgrade, Serbia (e-mail: bane@elfak.ni.ac.rs).

Dusan Milosević is with Faculty of Electrical Engineering, University of Nis, 14 Aleksandra Medvedeva, 18000 Belgrade, Serbia (e-mail: dusan.milosevic@elfak.ni.ac.rs).

Srdjan Ribar is with Faculty of Mechanical Engineering, University of Belgrade, Kraljice Marije 16, 11120 Belgrade Serbia, Belgrade, Serbia (e-mail: srdjanribar@gmail.com).

Ivana Radovic is with Institute of Nuclear Sciences 'VINČA', University of Belgrade, Mike Petrovica Alasa 12-14, 11000 Belgrade, Serbia (e-mail: ivanaradovic.80@gmail.com).

Markus Mohr is with University of Ulm, Institute of Functional Nanosystems, Ulm, Germany (e-mail: markus.mohr@uni-ulm.de).

Hans J. Fecht is with University of Ulm, Institute of Functional Nanosystems, Ulm, Germany (e-mail: hans.fecht@uni-ulm.de).

II. Graph Theory in Intergranular Relations

We have 3D structure of synthesized electronic material that should be presented or modelled using graph theory for calculation of parameters at grain boundary ([2]), shown on Fig.1. Let us consider 1D and 2D case, because this 2D case could be, later, easily generalized and mapped onto some 3D case.

A. Modelling with 1D and 2D graphs

In order to establish appropriate theoretical experiment, through which we will model this sintered plate with graph [2], lets simplify, and instead of 2D case (Fig.2), we will take one of 1D cases (Fig.3):

Fig. 1. Structure of grains (3D case)

Fig. 2. Structure of grains according to 2D case

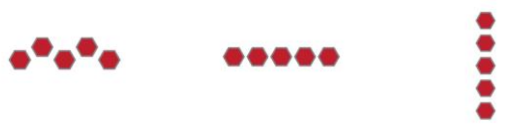


Fig. 3. Structure of grains according to 1D case

Now, let us make graph from one of those cases, for example the last one (vertical arrangement) with n=5 grains. Our goal is to get overall results in change of capacitance, in relation with those (measured) overall results. So, we have graphs, shown on Fig.4 (see [5]).

If we use experimental results from [5] and divide overall capacitance change to a local measure, between two grains, for example, for n=5 and for DC bias 25V, we have overall capacity change 69.45. This means that between neighboring vertices in graph we have capacity change $\frac{\Delta C}{C}$ of

$$\frac{\Delta C}{C} = 69.45 / (n-1) = 69.45 / 4 = 17.3625$$

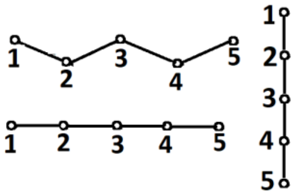


Fig. 4. Modelling with 1D graphs

Appropriate matrix of this graph (for n=5) is

$$A = \begin{bmatrix} 0 & 17.3625 & 0 & 0 & 0 \\ 17.3625 & 0 & 17.3625 & 0 & 0 \\ 0 & 17.3625 & 0 & 17.3625 & 0 \\ 0 & 0 & 17.3625 & 0 & 17.3625 \\ 0 & 0 & 0 & 17.3625 & 0 \end{bmatrix}$$

In that case, we have parallel connection of capacity changes, and overall capacity change is equal to 68.73 (see [10]).

For results in same experiment (see [5]), for n=5 and for DC bias 95V, we have overall capacity change 100.97. This means that between neighboring vertices in graph we have capacity change $\frac{\Delta C}{C}$ of

$$\frac{\Delta C}{C} = 100.97/(n-1) = 100.97/4 = 25.2425$$

Appropriate neighboring matrix of this graph (for n=5) is

$$A = \begin{bmatrix} 0 & 25.2425 & 0 & 0 & 0 \\ 25.2425 & 0 & 25.2425 & 0 & 0 \\ 0 & 25.2425 & 0 & 25.2425 & 0 \\ 0 & 0 & 25.2425 & 0 & 25.2425 \\ 0 & 0 & 0 & 25.2425 & 0 \end{bmatrix}$$

In that case, we have parallel connection of capacity changes, and overall capacity change is equal to 100.97 (see [5]).

Let us make graph from one of 2D cases, for example with n=8 graines (see [5], see Fig.5). So, we have graph with same number of vertices (n=8),

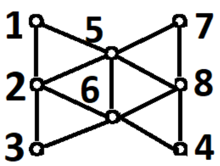


Fig. 5. Modelling with 2D graphs

If we use experimental results from [5] and if we divide overall capacitance change to local measure that characterizes relation between two grains (two vertices in graph). For example, for DC bias of 25V, we have overall capacity change of 69.45. This means that for *n=8*, *m=13* and for between neighboring vertices in graph we have capacity change $\frac{\Delta C}{C}$ of

$$\frac{\Delta C}{C} = 69.45/(n-1) = 69.45/12 = 5.7875$$

Appropriate neighboring matrix of this graph (see [10]) is

$$A = \begin{bmatrix} 0 & 5.7875 & 0 & 0 & 5.7875 & 0 & 0 & 0 \\ 5.7875 & 0 & 5.7875 & 0 & 5.7875 & 5.7875 & 0 & 0 \\ 0 & 5.7875 & 0 & 0 & 0 & 5.7875 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5.7875 & 0 & 5.7875 \\ 5.7875 & 5.7875 & 0 & 0 & 0 & 5.7875 & 5.7875 & 5.7875 \\ 0 & 5.7875 & 5.7875 & 5.7875 & 5.7875 & 0 & 0 & 5.7875 \\ 0 & 0 & 0 & 0 & 5.7875 & 0 & 5.7875 & 0 \\ 0 & 0 & 0 & 5.7875 & 5.7875 & 5.7875 & 5.7875 & 0 \end{bmatrix}$$

This results to overall capacitance change equal to 68.73.

B. *Modelling with 3D graphs*

We will now apply 3D graphs on calculation of breakdown voltage on $BaTiO_3$ sample (see [1]) with some predefined constraints. Relation between grains in sample is established and described and now we have mathematical approach for calculation of breakdown voltage using experimental results. As a result, we introduced mapping between property of sample and grain structure, then between grain structure and mathematical graph, using various crystal structures. The main idea was to apply 3D-graph theory for distribution of electronic parameters between the neighboring grains. We will use simple $BaTiO_3$ sample, shown on Fig.6.

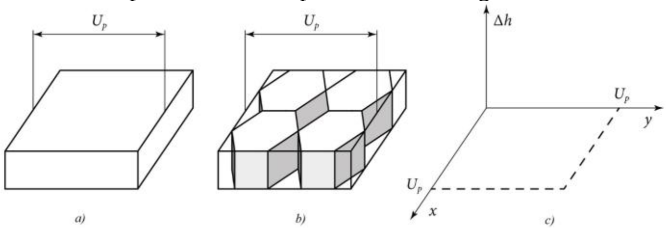


Fig. 6. $BaTiO_3$ sample with 4 grains.

Breakdown voltage U_p is obtained by measuring on whole sample, i.e.it is measured on two opposite sides of this sample. It is same in both directions, along x-axis and along y-axis (Fig. 2). Third dimension will be denoted with Δh, and that is vertical size of this sample, distance between top and bottom of the sample (tom and bottom of the grain).

If we do mapping of this grain sample (see [1]), onto 3D graph with eight vertices, shown on Fig.7. We put vertices on the top of each grain and at the bottom of each grain, assuming that there is also some value of U_p between grain ends. Breakdown voltage is very small and it can be neglected, leaving $U_p = \delta$, $(\delta \approx 0)$. Assume that "dimensions" of this problem are now $x = 2$, $y = 2$ and $\Delta h = 2$.

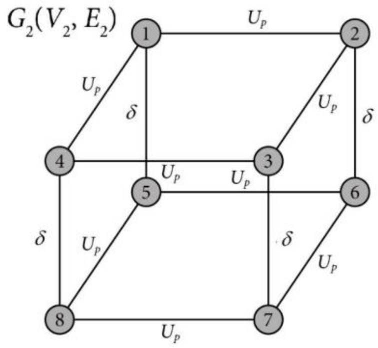


Fig. 7. 3D graph model of 2x2 grain sample, with 8 vertices.

Graph $G_2 = (V_2, E_2)$, has 8 vertices and 12 edges. Corresponding weight matrix of this graph is:

$$W_2 = \begin{bmatrix} 0 & U_p & 0 & U_p & \delta & 0 & 0 & 0 \\ U_p & 0 & U_p & 0 & 0 & \delta & 0 & 0 \\ 0 & U_p & 0 & U_p & 0 & 0 & \delta & 0 \\ U_p & 0 & U_p & 0 & 0 & 0 & 0 & \delta \\ \delta & 0 & 0 & 0 & 0 & U_p & 0 & U_p \\ 0 & \delta & 0 & 0 & U_p & 0 & U_p & 0 \\ 0 & 0 & \delta & 0 & 0 & U_p & 0 & U_p \\ 0 & 0 & 0 & \delta & U_p & 0 & U_p & 0 \end{bmatrix}$$

Each "horizontal" (along x-axis and along y-axis) edge of this graph has the same weight U_p, and each "vertical" edge (between top and bottom of the grain) has the same weight δ. In the case obtained from the experiment in [1], case I-16 $BaTiO_3$ ceramics without additives, we have:

$$W_2 = \begin{bmatrix} 0 & 12.54 & 0 & 12.54 & \delta & 0 & 0 & 0 \\ 12.54 & 0 & 12.54 & 0 & 0 & \delta & 0 & 0 \\ 0 & 12.54 & 0 & 12.54 & 0 & 0 & \delta & 0 \\ 12.54 & 0 & 12.54 & 0 & 0 & 0 & 0 & \delta \\ \delta & 0 & 0 & 0 & 0 & 12.54 & 0 & 12.54 \\ 0 & \delta & 0 & 0 & 12.54 & 0 & 12.54 & 0 \\ 0 & 0 & \delta & 0 & 0 & 12.54 & 0 & 12.54 \\ 0 & 0 & 0 & \delta & 12.54 & 0 & 12.54 & 0 \end{bmatrix}, (\delta \approx 0).$$

After that, we do the mapping of the same grain sample onto 3D graph with nine vertices, shown on next Fig.8. We put vertices on "top" of each grain and also on "bottom" of each grain (see [1]), like in previous case, breakdown voltage between those ends is very small, so we will assume that it is δ. New, ninth vertex is, by BCC principle (body-centered cubic crystal system imaginary point in the middle of this sample, and its appearance rises gives dimension of this problem from 2 to 3. Assume that "dimensions" of problem are $x = 2, y = 2$ and $\Delta h = 3$.

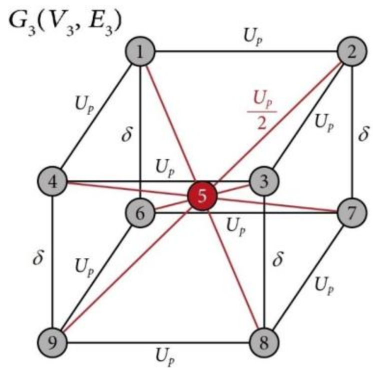


Fig. 8. 3D graph model (Body-centered in crystal system) with 9 vertices.

Graph $G_3 = (V_3, E_3)$, has a set of 9 vertices and a set of 20 edges (see [1]). Corresponding weight matrix of this graph is:

$$W_3 = \begin{bmatrix} 0 & U_p & 0 & U_p & U_p/2 & \delta & 0 & 0 & 0 \\ U_p & 0 & U_p & 0 & U_p/2 & 0 & \delta & 0 & 0 \\ 0 & U_p & 0 & U_p & U_p/2 & 0 & 0 & \delta & 0 \\ U_p & 0 & U_p & 0 & U_p/2 & 0 & 0 & 0 & \delta \\ U_p/2 & U_p/2 & U_p/2 & U_p/2 & U_p/2 & U_p/2 & U_p/2 & U_p/2 & U_p/2 \\ \delta & 0 & 0 & 0 & U_p/2 & U_p & 0 & U_p & 0 \\ 0 & \delta & 0 & 0 & U_p/2 & 0 & U_p & 0 & U_p \\ 0 & 0 & \delta & 0 & U_p/2 & U_p & 0 & U_p & 0 \\ 0 & 0 & 0 & \delta & U_p/2 & 0 & U_p & 0 & U_p \end{bmatrix}$$

Each "horizontal" edge of this graph has the same weight U_p, and each "vertical" edge, that goes "through" the grain have same weight δ. In case obtained from the experiment (see [1]), we have:

$$W_3 = \begin{bmatrix} 0 & 12.54 & 0 & 12.54 & 6.27 & \delta & 0 & 0 & 0 \\ 12.54 & 0 & 12.54 & 0 & 6.27 & 0 & \delta & 0 & 0 \\ 0 & 12.54 & 0 & 12.54 & 6.27 & 0 & 0 & \delta & 0 \\ 12.54 & 0 & 12.54 & 0 & 6.27 & 0 & 0 & 0 & \delta \\ 6.27 & 6.27 & 6.27 & 6.27 & 6.27 & 6.27 & 6.27 & 6.27 & 6.27 \\ \delta & 0 & 0 & 0 & 6.27 & 12.54 & 0 & 12.54 & 0 \\ 0 & \delta & 0 & 0 & 6.27 & 0 & 12.54 & 0 & 12.54 \\ 0 & 0 & \delta & 0 & 6.27 & 12.54 & 0 & 12.54 & 0 \\ 0 & 0 & 0 & \delta & 6.27 & 0 & 12.54 & 0 & 12.54 \end{bmatrix}, (\delta \approx 0).$$

Now, we can discuss same grain sample, mapped onto 3D graph, but now with twelve vertices, shown on Fig.9. New imaginary vertices are pointed in the middle of each surface of this sample, except upper surface and bottom surface. "Dimensions" of problem are $x = 2, y = 2$ and $\Delta h = 3$.

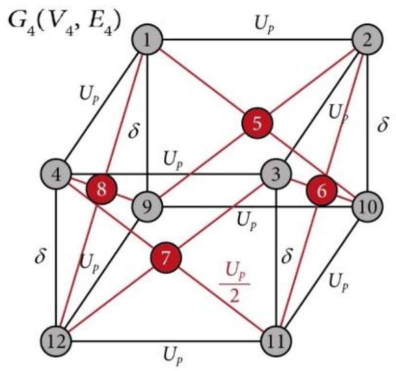


Fig. 9. 3D graph model (Face-centered in crystal system) with 12 vertices.

Corresponding weight matrix (see [1]) of this graph is

$$W_4 = \begin{bmatrix} 0 & U_p & 0 & U_p & U_p/2 & 0 & 0 & U_p/2 & \delta & 0 & 0 & 0 \\ U_p & 0 & U_p & 0 & U_p/2 & U_p/2 & 0 & 0 & 0 & \delta & 0 & 0 \\ 0 & U_p & 0 & U_p & 0 & U_p/2 & U_p/2 & 0 & 0 & 0 & \delta & 0 \\ U_p & 0 & U_p & 0 & 0 & 0 & U_p/2 & U_p/2 & 0 & 0 & 0 & \delta \\ U_p/2 & U_p/2 & 0 & 0 & 0 & 0 & 0 & 0 & U_p/2 & U_p/2 & 0 & 0 \\ 0 & U_p/2 & U_p/2 & 0 & 0 & 0 & 0 & 0 & 0 & U_p/2 & U_p/2 & 0 \\ 0 & 0 & U_p/2 & U_p/2 & 0 & 0 & 0 & 0 & 0 & 0 & U_p/2 & U_p/2 \\ U_p/2 & 0 & 0 & U_p/2 & 0 & 0 & 0 & 0 & U_p/2 & 0 & 0 & U_p/2 \\ \delta & 0 & 0 & 0 & U_p/2 & 0 & 0 & U_p/2 & 0 & U_p & 0 & U_p \\ 0 & \delta & 0 & 0 & U_p/2 & U_p/2 & 0 & 0 & U_p & 0 & U_p & 0 \\ 0 & 0 & \delta & 0 & 0 & U_p/2 & U_p/2 & 0 & 0 & U_p & 0 & U_p \\ 0 & 0 & 0 & \delta & 0 & 0 & U_p/2 & U_p/2 & U_p & 0 & U_p & 0 \end{bmatrix}$$

Each "horizontal" edge of this graph has same weight U_p, and each "vertical" edge, that goes "through" the grain have same weight δ. For $BaTiO_3$ ceramics (see [1]) without additives, we have:

$$W_4 = \begin{bmatrix} 0 & 12.54 & 0 & 12.54 & 6.27 & 0 & 0 & 6.27 & \delta & 0 & 0 & 0 \\ 12.54 & 0 & 12.54 & 0 & 6.27 & 6.27 & 0 & 0 & 0 & \delta & 0 & 0 \\ 0 & 12.54 & 0 & 12.54 & 0 & 6.27 & 6.27 & 0 & 0 & 0 & \delta & 0 \\ 12.54 & 0 & 12.54 & 0 & 0 & 0 & 6.27 & 6.27 & 0 & 0 & 0 & \delta \\ 6.27 & 6.27 & 0 & 0 & 0 & 0 & 0 & 0 & 6.27 & 6.27 & 0 & 0 \\ 0 & 6.27 & 6.27 & 0 & 0 & 0 & 0 & 0 & 0 & 6.27 & 6.27 & 0 \\ 0 & 0 & 6.27 & 6.27 & 0 & 0 & 0 & 0 & 0 & 0 & 6.27 & 6.27 \\ 6.27 & 0 & 0 & 6.27 & 0 & 0 & 0 & 0 & 6.27 & 0 & 0 & 6.27 \\ \delta & 0 & 0 & 0 & 6.27 & 0 & 0 & 6.27 & 0 & 12.54 & 0 & 12.54 \\ 0 & \delta & 0 & 0 & 6.27 & 6.27 & 0 & 0 & 12.54 & 0 & 12.54 & 0 \\ 0 & 0 & \delta & 0 & 0 & 6.27 & 6.27 & 0 & 0 & 12.54 & 0 & 12.54 \\ 0 & 0 & 0 & \delta & 0 & 0 & 6.27 & 6.27 & 12.54 & 0 & 12.54 & 0 \end{bmatrix},$$

where $\delta \approx 0$.

III. SINTETIZED DIAMONDS AS BIOMIMETIC MATERIALS

Materials science is spreading into all fields of basic and applied science. One of the biggest influence on everydays life in bioceramics materials is the application of these substrates for medical engineering (see [4]). The most prominent bio-relevant properties of a biomaterial are chemical inertness and bio-durability, so submicro diamond is coming into the center of our research interest. This diamond structure is an unique biomaterial with specific compatible characteristics, which are not exactly only one based on its bio-chemical properties. Physical-chemical biocompatibility directions, including data from nano interface water layers could be defined on nanodiamond substrates (see [4]). If we, theoretical, split diamond surface into small particles, we can implement similar approach like with grains in [5]. If we take a small part of diamond surface, of "dimension" 2x2x1 (Fig.6), we can detect and assign graph vertices in upper surface of diamond, shown on Fig.10.

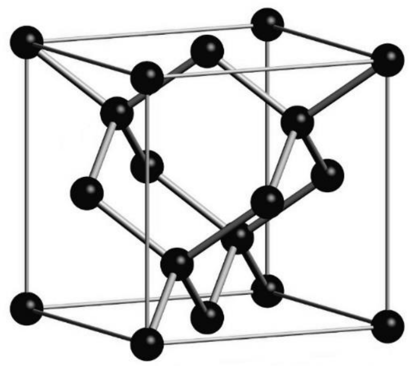


Fig. 10. Syntetized diamond internal structure.

Appropriate graph is shown on Fig.11. We can assume some property (any electrical, dielectrical or magnetic property), noted with v, and we can assign values on those graph edges, we can represent this graph (see [4]) with weight matrix W.

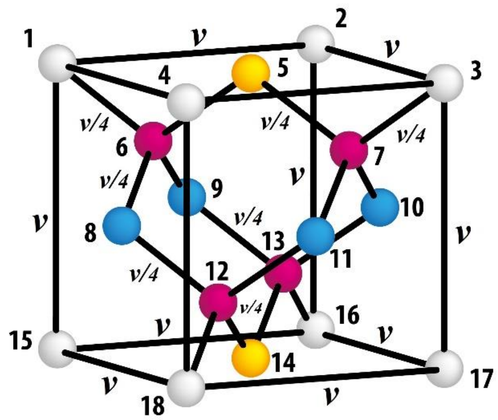


Fig. 11. 3D graph model of syntetized diamond

Weight matrix of this graph (see [4]) is:

$$W = \begin{bmatrix} 0 & v & 0 & v & 0 & v/4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & v & 0 & 0 & 0 \\ v & 0 & v & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & v & 0 & 0 \\ 0 & v & 0 & v & 0 & 0 & v/4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & v & 0 \\ v & 0 & v & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & v \\ 0 & 0 & 0 & 0 & 0 & v/4 & v/4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ v/4 & 0 & 0 & 0 & v/4 & 0 & 0 & v/4 & v/4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & v/4 & 0 & v/4 & 0 & 0 & 0 & 0 & v/4 & v/4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & v/4 & 0 & 0 & 0 & 0 & 0 & v/4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & v/4 & 0 & 0 & 0 & 0 & 0 & 0 & v/4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & v/4 & 0 & 0 & 0 & 0 & 0 & v/4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & v/4 & 0 & 0 & 0 & 0 & v/4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & v/4 & 0 & 0 & v/4 & 0 & 0 & v/4 & 0 & 0 & 0 & v/4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & v/4 & v/4 & 0 & 0 & 0 & v/4 & 0 & v/4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & v/4 & v/4 & 0 & 0 & 0 & 0 & 0 \\ v & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & v & 0 & v \\ 0 & v & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & v/4 & 0 & v & 0 & v & 0 \\ 0 & 0 & v & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & v & 0 & v \\ 0 & 0 & 0 & v & 0 & 0 & 0 & 0 & 0 & 0 & 0 & v/4 & 0 & 0 & v & 0 & v & 0 \end{bmatrix}$$

Each “horizontal” edge of this graph has weight v, direct “vertical” edge also has weight v, and “skew” edges, through structure has weight $v/4$. Why divided with 4? Because in this model every algebraic path through this graph, starting from upper surface down to lower surface, consist of 4 edges, length of algebraic path is 4 (see [4]).

The ultra-nanocrystalline synthetized diamonds are very advanced materials for biomedical and other applications. We pointed out to a complex relation between graph theory and electrophysical parameters of the consolidated nano-diamonds is presented. We performed and explained related experimental procedure with results data (see [4]). By this method we provided way for defining the electrophysical parameters on micro and nano level of grains and pores, what is important for further designing microelectronic structures and advance miniaturization.

IV. Outlook

In the future research, it is possible to correlate graph theory application on some other ceramic materials characteristics and to use it for some other values calculation on microlevel, for example electrical conductivity, thermal conductivity, etc. Also, it could be very important to compare the results based on neural networks and graph application, simultaneously. The nanocrystalline synthetized diamonds are very advanced materials for high-tech applications especially in medicine, electronic and space research. Further application of the graph theory decelopment could make some new and unexpected contributions in this field.

V. Conclusion

In this research paper we intended to make an overview on a complex relation between graph theory, BaTiO3-ceramics and synthetized diamonds electrophysical parameters calculations, and also designed microeletronic intergranular relations. Before the graph and neural network theories (see [10-14]), we have used different experimental methods for global measurements and collecting the physical parameters from the sample surfaces, based mostly on the statistic distribution.

Now we can define values directly at the level of the grains and pores, and, also, between them. Based on this approacch we provided a way for defining other parameters on micro and nano grains and pores constituents, what is important for advance predicting microelectronic structures and related parameters, in electronic material sciences in general.

References

[1] V.V.Mitic, B.Randjelovic, I.Ilic, S.Ribar, **C**.An-Lu, A.Stajcic, B.Vlahovic.“The 3D graph approach for breakdown voltage calculation in BaTiO3 ceramics”, IJMPB, Vol 35 (7) (2021) pp. 2150103.

[2] V.V.Mitic, G.Lazovic, B.Randjelovic, V.Paunovic, I.Radovic, A.Stajcic, B.Vlahovic: “Graph Theory Applied to Microelectronic Intergranular Relations”, Ferroelectrics, 570 (2021) p. 145-152.

[3] V. Mitić, C.A.Lu, B.Randjelović, V. Paunović, I.Radovic, A. Stajcic, J.Kotnik, H.Fecht, B. Vlahovic, “3D graph theory application on modified nano BaTiO3 electronic ceramics”, Int.Conf. Electronic Materials and Applications EMA-2021, Virtual, January 19-21, 2021.

[4] B.Ranđelović, V.V.Mitic, S.Ribar, M.Cebela, M.Mohr, H.Fecht, B.Vlahovic: “Graph Theory Approach in Synthetized Diamond Electrophysical Parameters Defining”, in: Bioceramics, Biomimetic and other Compatible Materials Features for Medical Applications (Ed: S.Najman., V.V.Mitic, T.Groth, M.Barbeck, P.Yu Chen, Z.Sun), Springer Nature, Cham, Switzerland, 2021, accepted.

[5] B.Randjelovic, V.V.Mitic, S.Ribar, I.Radovic, A.Stajcic, I.Novakovic, B.Vlahovic. “Ceramics, Materials, Microelectronics and Graph Theory new Frontiers”, MPLB, Vol 34 (34) (2020) pp.2150159.

[6] V. Mitlić, G. Lazović, B.M. Randjelović, V. Paunović, J.M. Wu D.Mancić,, J.R. Hwu, “Fractal Microeletronic Frontiers and Graph Theory Applications”, Int.Conference MS&T **2019**, Book of Abstracts, Portland, USA, 29 September – 03 October 2019.

[7] V. Mitić, G. Lazović, C.A.Lu, I.Radovic, V. Paunović, A. Stajcic, B.Randjelović, S.Ribar, B. Vlahovic, “Investigation of intergranular dielectric properties within the relation between fractal, graph and neural networks theories”, Int.Conf. Electronic Materials and Applications EMA-2021, Virtual, January 19-21, 2021.

[8] V. Mitić, C. Serpa, C.A.Lu, V. Paunović, I.Radovic, B.Randjelović, S.Ribar, A. Stajcic, B. Vlahovic, “Fractal, graph and neural network theories applied on BaTiO3 electronic ceramics”, Int.Conf. Electronic Materials and Applications EMA-2021, Virtual, January 19-21, 2021.

[9] V. Mitić, I.Radovic, C.A.Lu, S.Ribar, B.Randjelović, A. Stajcic, H. Fecht, B. Vlahovic, “Neural networks and applied graph theory approaches for intergranular properties measurements investigation”, Int.Conf. Electronic Materials and Applications EMA-2021, Virtual, January 19-21, 2021.

[10] Mitic V. V.. *et al*. The Artificial Neural Networks Applied for Microelectronics Intergranular 323 Relations Determination. *Integrated Ferroelectics*. 212 .(2020).

[11] V.V.Mitic, S.Ribar, B.Randjelovic, C.-A. Lu, I.Radovic, A.Stajcic, I.Novakovic, B.Vlahovic, Neural Networks and Microelectronic Parameters Distribution Measurements depending on Sinthering Temperature and Applied Voltage, MPL B, 34 (**35)** (2020) pp. 2150172.

[12] Mitic V. V *et al.* Sinterign Temperature Influenceo on Grains Function Distribution by Neural Network Application, Thermal Science (2021) (accepted).

[13] Ribar S. *et al.* Neural Networks from Biophysical Applications in Microelectronics Parameters 336 Measurements. in: *Bioceramics. Biomimetic and other Compatible Materials Features for* 337 *Medical Applications.* (Ed. Najman S.. Mitic V.. Groth T.. Barbeck M.. Yu Chen P.. Sun Z.). 338 Springer Nature. Cham. Switzerland. 2021. accepted for publication

[14] Ribar S. *et al*. Neural Networks Application on Human Skin Biophysical Impedance 319 Characterisations. *Biophysical Reviews & Letters*. 16.(2021). pp. 2021 – 19

Biomolecules and Brownian Motion

Vojislav V. Mitić, Bojana Marković, Sanja Aleksić, Dušan Milošević, Branislav Randjelović, Ivana Ilić, Jelena Manojlović and Branislav Vlahović

***Abstract*—Structures and different life functions of microorganisms, like motion, are based on molecular biology processes, which comprise molecular and submolecular particles. It is very important to determine relation between molecular and microorganisms levels. The aim of our research is the analysis of Brownian motion as a general phenomenon and the consequence of structures hierarchy from molecular to microorganisms level. If we approach this idea from the aspect of biomimetic correlations at the level of the alive and nonalive matter system particles, the condensed matter particles could be considered as a part of micro, nano and molecular microorganisms structures. In this research we used the experimental results of bacterial motion influenced by different energy impulses. The important goal of this research paper is to obtain significant data regarding Brownian motion in the frame of fractal nature similarities, as an integrative property of living and nonliving systems particles processes. This opens new frontiers for submicroelectronics relations within the integrated supermicro biophysical systems. This is a potential new trend in nowadays advanced research, where we integrate the knowledges of complex relations between the electrons or other particles and their clusters as joint structures in alive and condensed matter, what could be a possible direction for new microelectronics complex biodevices and integrations.**

***Index Terms*—molecular biology; microorganisms; Brownian motion; biomimetic; fractals.**

Vojislav V. Mitić is with the Faculty of Electronic Engineering, University of Nis, 14 Aleksandra Medvedeva, 18000 Nis, Serbia and Institute of Technical Sciences of SASA, 35 Kneza Mihaila, 11000 Belgrade, Serbia (e-mail: vmitic.d2480@gmail.com).

Bojana Marković is with the Faculty of Electronic Engineering, University of Nis, 14 Aleksandra Medvedeva, 18000 Nis, Serbia (e-mail: bojana.markovic123@gmail.com).

Sanja Aleksić is with the Faculty of Electronic Engineering, University of Nis, 14 Aleksandra Medvedeva, 18000 Nis, Serbia (e-mail: sanja.aleksic@elfak.ni.ac.rs).

Dušan Milošević is with the Faculty of Electronic Engineering, University of Nis, 14 Aleksandra Medvedeva, 18000 Nis, Serbia (e-mail: dusan.milosevic@elfak.ni.ac.rs).

Branislav Randjelović is with the Faculty of Electronic Engineering, University of Nis, 14 Aleksandra Medvedeva, 18000 Nis, Serbia and Faculty of Teachers Educationin Prizren - Leposavić, University of Priština - Kosovska Mitrovica, Nemanjina, 38218 Leposavić, Serbia (e-mail: branislav.randjelovic@elfak.ni.ac.rs).

Ivana Ilić is with the Medical Faculty, Department for Mathematics and Informatics, University of Nis, 81 Bul. Dr Zorana Djindjića, 18000 Nis, Serbia (e-mail: ivanailicc3@gmail.com).

Jelena Manojlović is with the Faculty of Mechanical Engineering, University of Nis, 14 Aleksandra Medvedeva, 18000 Nis, Serbia (e-mail: jmanojlo@gmail.com).

Branislav Vlahović is with the North Carolina Central University, USA Durham NC, USA (e-mail: vlahovic@nccu.edu).

I. Introduction

Phenomena that exist in both living and nonliving matter, as electron and molecular motion, are to be considered from different aspects and they require a multidisciplinary approach. Molecular biology, as a life science, studies various processes in alive organisms at the molecular level. Thus, electron motion in biomolecules as well as molecular motion, belong within the molecular biology and molecular bioelectronics research field.

Electron motion is based on the same principles in alive organisms and in condensed matter, so it should be observed as a joint property to get the complete insight into this fundamental process. Based on our current knowledge and possibilities, we are not able to determine electron motion. However, we can determine motion of the molecules that contain those electrons. Every molecule which is moving, carries an electron cluster making it move as well. Within molecules as parts of biosystems, existing atoms and electrons „are not aware of" whether they are a part of an alive organism or condensed matter. That is very important because it distinguishes the molecule as the significant integrative factor between living and nonliving systems.

The idea of biomimetic correlation between molecular and submolecular particles of alive and nonalive matter, in the frame of the Brownian motion fractal nature characterization, implies the entire new possibility of considering these particles motion as the biunivocal phenomenon. Microorganisms and viruses demonstrate biomimetic similarities with condensed matter particles motion because of their dimensions and motion patterns [1]. Bacteria motility behavior, which implies velocity, direction and trajectory, is also influenced by environmental changes like temperature, pH, or different energetic impulses [2].

We used Brownian motion fractal nature, as a general characteristic of both alive and nonalive systems, to establish the relation between these two different systems, but consisting of identical particles. Also, we plan to develop the biophysical-mathematical asymptotic approaching model for living organisms and condensed matter particles, as they are biunivocaly correspondent.

One of the objectives of our research is to explain the Brownian motion as the joint characteristic of biomolecule and physical system particles. Controling particles motion and predicting their trajectories [3] in this kind of biomimetic approach [4], provides new perspectives for further microelectronics miniaturization and electronics parameter integrations [5,6].

So, our goal is to characterize the molecular motion based on Brownian motion within the combination of the experimental and the theoretical experiment results.

II. Experimental Methods and Procedures

We performed real experiments with diverse bacteria in a liquid phase under the influence of variant energy impulses [7], among which were some different music impulses (Figure 1). In that sense, we got some data from which we generated the two and three dimensional diagrams [2,7].

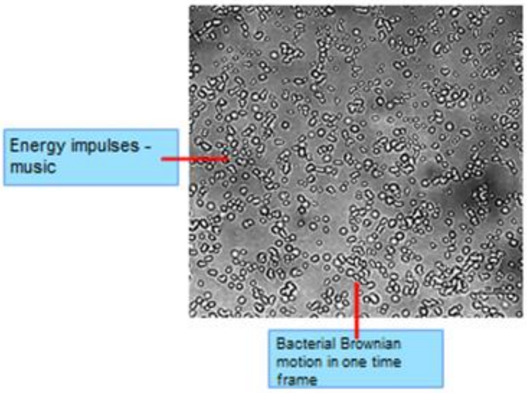


Fig. 1. The bacterial motion experiment diagram.

Now, our intension is to generate the analytical forms. Also, we used some other available research results which are based on molecular motions [8]. In this case we treat the molecules like clusters of electrones in different matter organizations from atomic to molecular level. This way we avoid the lack of worldwide research recorded electron motions. Hence, we now observe the molecules as “packages“ of electrons or other particles. From the other hand, we also analyze all of these molecules as a part of alive bacterial matter. At the end, we can jointly understand the biophysical integrated systems with one important characteristic and that is just the Brownian motion.

A. *Mathematical Background*

It is well known that it is possible to create some mathematical model describing the relationship between several quantities. By using one of them, called the dependent variable (we will denote it by y) and one or more quantities, called the independent variables (we will denote them by $x_1, x_2 \dots x_n$,), we can obtain a model which represents a linear relationship between the dependent variable and independent variables in the form:

$$y = a_1 x_1 + a_2 x_2 + \dots + a_n x_n + b, \tag{1}$$

where $a_1, a_2 \dots a_n,\ b$ are real (or complex) numbers. If the value of the variable y depends only on one independent variable x, then (1) has a form:

$$y = ax + b. \tag{2}$$

This formula could be obtained by applying various numerical approximation models. One of very useful approximation models is the least square approximation.

The least squares method (also called discrete mean square approximation [9]) belongs to the so-called best approximations, i.e. approximation methods in which the criterion is the minimization of the error according to one of the norms. Specifically, this is the norm L^2, ie. the total sum of the squares of the errors in the approximation nodes is minimized [9,10]. Some of the interesting applications of those methods are already given in [11-13].

B. *Main Results*

We applied the technique to the data in Table I, which describes the coordinates of bacteria locations, during movement through coordinate system [7].

TABLE I
Bacteria locations coordinates

i	x_i	y_i	z_i
1	0	0	0
2	0.1043	-0.3698	-0.2869
3	0.0521	-0.4622	-0.3641
4	0.0521	-0.2773	-0.4809
5	0.0521	-0.2773	-0.7842
6	0.0521	-0.1849	-0.7605
7	0.1564	-0.5547	-0.7709
8	0.2607	-0.7396	-0.7757
9	0.5213	-0.7396	-1.0163
10	0.4170	-0.8320	-0.9330
11	0.3649	-0.8320	-0.9349

Based on the data from Table I, we obtained the 3D diagram presented in Figure 2.

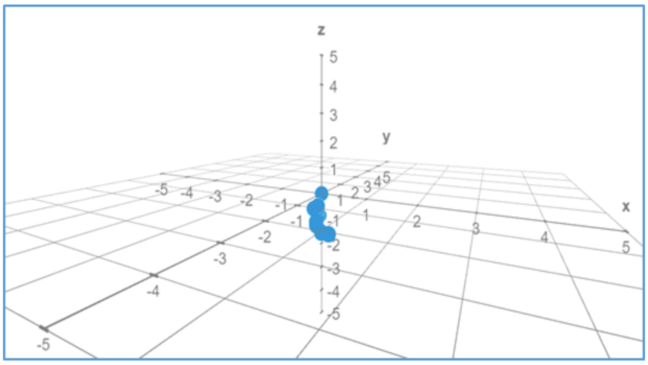


Fig. 2. The points of the bacteria locations in 3D.

Next, we considered four molecule location points in 3D given in Table II and presented in Figure 3.

TABLE II
Molecule locations coordinates

i	x_i	y_i	z_i

1	2	5.8	4
2	2.2	2	4.2
3	2.5	4.4	4.5
4	2.8	3.2	5.2

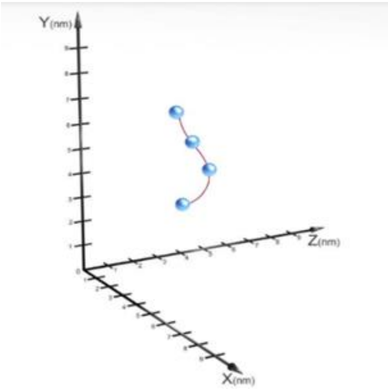


Fig. 3. 3D diagram of molecule motion in different time intervals.

III. RESULTS AND DISCUSSION

We applied multiple linear regression, to determine the mutual dependence of the coordinates and to obtain explicit formula for predicting and calculating positions.

Based on the data from Table I, we will apply the procedure of forming an approximation function

$$\varphi(x, y) = ax + by + c, \tag{3}$$

by using the least squares method. Thus, by applying the least square approximation (the operational software was statistical package in Excel) on the given data sets, we obtained next results considering the best linear fit for the presented model: the coefficients of the resulting linear function are respectfully a=–1,47999912, b=–0,02844679 and c=–0,36733904 and the estimated regression function is of the form:

$$\varphi(x, y) = -0.746871x - 0.421536y - 0.306160 \tag{4}$$

We can compare values and precision of dependent variables z_i in given points and results obtained by formula trough the absolute and relative error (Table 3).

TABLE III
COMPARISON BETWEEN REAL AND APPROXIMATE COORDINATES, ABSOLUTE AND RELATIVE ERROR

x_i	y_i	z_i	φ_i	Δ	%
0	0	0	-0.3061600	0.3061600	
0.1043	-0.3698	-0.2869	-0.5399426	0.2530427	-88.20%
0.0521	-0.4622	-0.3641	-0.5399059	0.1758059	-48.29%
0.0521	-0.2773	-0.4809	-0.4619639	0.0189361	-3.94%
0.0521	-0.2773	-0.7842	-0.4619639	0.3222361	-41.09%
0.0521	-0.1849	-0.7605	-0.4230139	0.3374860	-44.38%
0.1564	-0.5547	-0.7709	-0.6567966	0.1141034	-14.80%
0.2607	-0.7396	-0.7757	-0.8126373	0.0369373	-4.76%
0.5213	-0.7396	-1.0163	-1.0072718	0.0090281	-0.89%
0.4170	-0.8320	-0.9330	-0.9683232	0.0353232	-3.79%
0.3649	-0.8320	-0.9349	-0.9294112	0.0054888	-0.59%

The plot (4) obtained with the least squares method is presented in Figure 4.

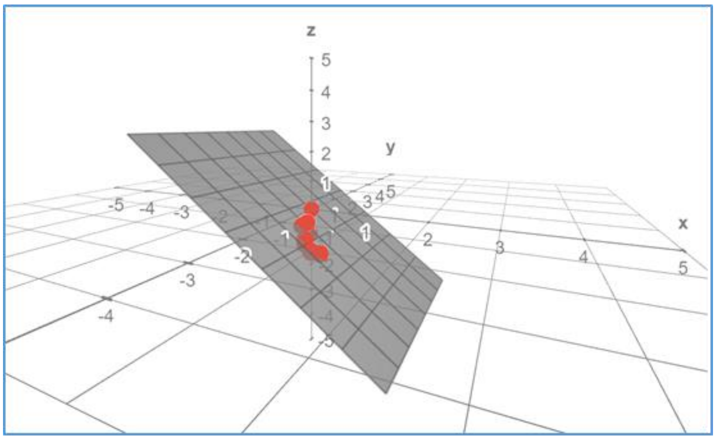


Fig. 4. The approximation plot with marked red points from Table III.

Similarly, as in previous procedure applied on bacterial motion experimental data, we obtained next results for molecule motion in different time intervals, considering the best linear fit for the presented model: the coefficients of the resulting linear function are respectfully a=1.4685067, b=0.0035386 and c=0.973673 and the estimated regression function is of the form:

$$\varphi = 1.4685067x + 0.0035386y + 0.973673. \tag{5}$$

Next, by using the estimated regression function (5) and by implementing the 2D coordinates we obtained the estimated dependent values of the z-coordinates, presented in the Table IV, together with the evaluated absolute and relative error of this approximation:

TABLE IV
Z- COORDINATES WITH THE ABSOLUTE AND RELATIVE ERROR

x_i	y_i	z_i	φ_i	Δ	%
0	0	0			
2	5.8	4	3.9312102	0.0687898	1.72%
2.2	2	4.2	4.2114649	0.0114649	0.27%
2.5	4.4	4.5	4.6605095	0.1605096	3.57%
2.8	3.2	5.2	5.0968150	0.1031850	1.98%

The plot (5) obtained with the least squares method is presented in Figure 5.

If we observe alive and nonalive matter particles as a hierarchical phenomenon, we can consider an atom as a cluster of electrons and other particles, a molecule as a cluster of atoms with already mentioned particles which are penetrating each other from their orbitals in interatomic relations within the molecule, and a microorganism as a cluster of molecules. Nowadays fundamental research and science do not have high-tech and also resolution possibilities to recognize, separately, the electron motion. We can consider only the indirect effects. Here, we must stress the complexity in the matter based on quantum mechanical principles and Heisenberg uncertainty principle, as well, in all of these considerations.

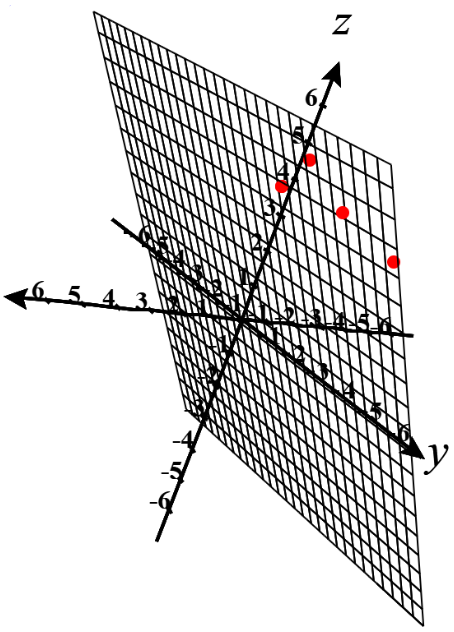


Fig. 5. The approximation plot with marked red points from Table IV.

Each bacterial cell comprises 2 – 4 millions of protein molecules [14], which implies that the total number of molecules per bacterial cell is much higher. This is just one comparison. We can observe the effect of electron motion at the molecular and microorganisms level. So definitely, the particles motion based on Brownian motion fractals effects, is the base for deeply understanding all of these processes within the submicro scale sizes with the joint characteristic which we can nominate as "actio in distans" in motion.

In this paper we introduced two mathematical analytical forms: one for bacterial and second for molecular motion, which are characterized by Brownian motion. In that sense we would like to establish a relation between these two mathematical analytical forms considering molecule number ratio. In this way we could determine asymptotic approaching of two mathematical functions towards fractals biomimetical self-similarity. This is the idea for our further research.

IV. Outlook

All of these results and innovative methods in mathematical analysis and discussions are just the original start in this field. In the next step we plan to develop a method by which we can establish the relation between the analytical mathematics and asymptotic approaching of alive and nonalive matter particles, with an idea to do it by fractal similarities.

V. Conclusion

In this paper we performed the results regarding molecular and bacterial motion. Regarding molecular motion, comprising electron motion, we provided mathematical analytical forms, in order to substantially characterize this motion. We analyzed bacterial motion influenced by different energy impulses, like music, and presented obtained data, showing random bacterial trajectories based on Brownian motion, in mathematical analytical forms, as well. Thus, in this stage we fulfilled and satisfied the idea of "electron clusters" in biomolecules which are also a part of microorganisms. Here, we define just a beginning of the joint integration in biophysical systems with general characteristic which is the Brownian motion.

Acknowledgment

The authors gratefully acknowledge the support of Ministry of Education, Science and Technological Development of Serbia, and The National Science Foundation of North Carolina, USA, for this research.

References

[1] V.V.Mitic,G.Lazovic, D.Milosevic, E.Ristanovic, D.Simeunovic,S.-C. Tsay, M.Milosevic and B.Vlahovic, "Brownian fractal nature Coronavirus motion," *Mod. Phys. Lett. B*, vol. 35, no. 04, pp. 2150076, Jan. 2021.

[2] V.V. Mitic, G. Lazovic , D. Milosevic, C.-A. Lu, J. Manojlovic, S.-C. Tsay, S. Kruchinin and B. Vlahovic, "Brownian motion and fractal nature," *Mod. Phys. Lett. B*, vol.34, no. 19n20, pp. 1-11, July, 2020.

[3] Z. B. Vosika, V. V. Mitić, G. Lazović, V. Paunović and Lj. Kocić, "Meso-kinetics of one time relaxation electrical processes in BaTiO3 ceramics—modified Boltzmann-Poisson model," *Ferroelectrics*, vol. 531, pp. 38–50, Nov, 2018.

[4] J. Mei, X. Peng, Q. Zhang, X. Zhang, T. Liao, V.V. Mitic and Z, Sun, "Bamboo-membrane inspired multilevel ultrafast interlayer ion transport for superior volumetric energy storage," *Adv. Funct. Mater*, May, 2021.

[5] V.V. Mitic, S.Ribar, B.Randjelovic, C.-A. Lu, I. Radovic, A. Stajcic, I.Novakovic and B.Vlahovic, "Neural networks and microelectronics parameters distribution measurements depending on sintering temperature and applied voltage," *Mod. Phys. Lett. B*, vol.34, no. 35, pp. 2150172, Dec, 2020.

[6] V. V. Mitic, G. Lazovic, S. Ribar, C.-A. Lu, I. Radovic, A. Stajcic, H. Fecht and B. Vlahovic, "The artificial neural networks applied for microelectronics intergranular relations determination," *Integr. Ferroelectr*, vol. 212, pp. 135–146, Nov, 2020.

[7] V.V. Mitic, G. Lazovic, D. Milosevic, J. Manojlovic, E. Ristanovic, D. Simeunovic, S.-C. Tsay, M. Milosevic, M. Sokovic and B. Vlahovic, "Brownian motion fractal nature frontiers within the matter," in *Biocompatible and Biomimetic Features of Materials for*

Biomedical Applications – New Insights, Switzerland, Springer, 2021, ch. 10.

[8] S. Bhunia, A. Rana, P. Roy, D. J. Martin,M. L. Pegis, B. Roy and A. Dey, "Rational design of mononuclear iron-porphyrins for facile and selective 4e-/4H+ O2 reduction: activation of O-O bond by 2nd sphere hydrogen bonding," *J. Am. Chem. Soc*, vol. 140, no.30, pp. 9444-9457, July, 2018.

[9] G.V.Milovanović, *Numerical analysis II*, 3rd ed. Belgrade, Serbia, Naucna Knjiga,1991, VIII+210 pp.

[10] G.V.Milovanović and M.A.Kovačević, *A collection of solutions for problems in numerical analysis*, 3rd ed. Belgrade, Serbia, Naucna Knjiga,1991, IV+278 pp.

[11] B.Ranđelović, I.Radojković and B.Gajić, "One Application of Discrete Mean Squere Approximation For Prediction of Voluntary Pension Fund Development," *Appl. Math. and Comp. Science*, vol 5, no. 1, University of Nis, Faculty of Science, pp.7-15, 2020.

[12] I.D.Radojkovic**,** B.Gajic and B.M.Randjelovic, "Predicting the Development of Voluntary Pension Funds in Serbia by Applying The Mathematical Method of Linear Regression," *J. of Economic and Social Development – Resilient Society*, vol. 8, no. 1, pp. 2-13, 2021.

[13] I.D.Radojkovic, B.M.Randjelovic and I.D.Ilic, "Modelling Development of Volontary Pension Fund Using Mathematical Model Of Approximation with Lagrange Interpolation Polynomials," *Ekonomske teme*. vol. 59, no.2, pp.243-257, 2021.

[14] R. Milo, "What is the total number of protein molecules per cell volume? A call to rethink some published values," BioEssays, vol.35, no. 12, pp. 1050-1055, Dec. 2013.

Reconstruction of fiber reinforcement in epoxy-based composite

Aleksandar Stajčić, Vojislav V. Mitić, Cristina Serpa, Branislav Randjelovic, Ivana Radović

Abstract—**Polymer matrix composites (PMCs) are very attractive materials due to a possibility to achieve versatile properties by combining with ceramic or metal reinforcement in different shapes and sizes. As a result, PMCs have found application in nearly every field, from household appliances to aerospace industry. Modern microelectronic devices contain conductive polymers with fillers that enhance their electrical properties. In addition, PMCs are being used as insulators and adhesives, contributing to the long life of electronic devices. Epoxy resins are the most commonly used insulators and adhesives. In order to improve their fracture toughness, glass fibers can be used as an efficient reinforcement. However, with the purpose of designing a composite with good mechanical properties and durability, deep knowledge of microstructure is required. In addition, microstructural analysis can be used to connect shape and size of pores or reinforcement with various physical properties. Fractal nature analysis is a valuable mathematical tool that can be employed for different shapes and forms rendering. In this manner, successful design and prediction of composite's properties could be obtained. In this research, field emission scanning electron microscopy (FESEM) images were used for fractal analysis of glass fibers, with the aim of reconstructing the shape.**

Index terms— **Fractal analysis; Composites; Epoxy; Microelectronics.**

I. Introduction

Composites represent multiphase materials containing two or more phases distinct by an interface [1-3]. Physical properties of composites are significantly different from the initial constituents' properties [4,5]. Civilization modernization was followed by increased production and development of composites both at the research and industrial scale. With composite design and processing, structural, thermal, electrical and other properties of the constituents are improved. Synergetic effect of matrix and reinforcement properties results in lightweight materials with high toughness and strength that can be resistant to corrosion, chemicals and temperature [6]. Nowadays, there are many efficient ways to process composite with targeted functional properties, with structure control at micro- and nanoscopic level [7-12]. Polymers are very attractive as matrix materials in composites, due to their low price and density, as well as wide range of good physical and chemical properties. Conductive polymers like poly(vinylidene fluoride) polypyrrole are investigated and implemented in various electronic parts. Composites with epoxy matrix are used as insulators, offering device stability and durability [13]. Since epoxy generally has low fracture toughness, fibers are usually incorporated for the reinforcement. They increase toughness, specific strength and modulus of elasticity [14,15] In order to avoid delamination and fiber separation from the matrix, thorough insight into the structure in desirable [16-18]. Fractal nature analysis is a powerful mathematical tool for the investigation of materials morphology that is used for characterization of grains and pores [19]. However, it can also serve to describe and predict the shape and size of reinforcement, enabling more efficient future composite design. In this manner, processing-structure-property circle can be closed. FESEM images can be used for the shape and size reconstruction of the fibers.

A. Fractal nature analysis

The fractal nature exists within physical systems structures and contact surfaces, from microstructures, down to the nano-scale level, up to the global bulk and massive shapes. Fractal nature analysis presents a possible approach for the investigation of contact phenomena establishing the grain contacts models, offering ceramics and other materials structure analysis, description and prediction of grains' and pores shape, along with relations between structure and electric-dielectric properties. Contribution of fractals correction could be observed and explained on intergranular Heywang capacity model, Schottky barrier, Curie-Weiss law, Clausius-Mossotti relation and other parameters in the field of dielectric and ferroelectric materials, by introducing complex fractal correction factor α (grain and pores surface influence and Brownian particle motion). The development of fractal

A. Stajčić is with the Center of Microelectronic Technologies, Institute of Chemistry, Technology and Metallurgy – National Institute of the Republic of Serbia, University of Belgrade, Belgrade, Serbia, stajcic@nanosys.ihtm.bg.ac.rs
V. Mitić is with the University of Nis, Faculty of Electronic Engineering, Nis, Serbia and Institute of Technical Sciences, Serbian Academy of Sciences and Arts, Belgrade, Serbia vmitic.d2480@gmail.com
C. Serpa is with the ISEL - Instituto Superior de Engenharia de Lisboa do Instituto Politécnico de Lisboa cristinaserpa@hotmail.com
B Ranđelović is with the University of Nis, Faculty of Electronic Engineering, Nis, Serbia and University of K. Mitrovica, Faculty of Teachers Education, Leposavic, Serbia bane@elfak.ni.ac.rs
I. Radović is with the Department of Physical Chemistry, 'VINČA' Institute of Nuclear Sciences - National Institute of the Republic of Serbia, University of Belgrade, Belgrade, Serbia ivana_r@vinca.rs

analysis idea is inspired by self-similarity in nature biosystems, where the chaotic structures could be controlled by recognized geometry structure or just to have disorder controlled towards the order. One of the goals is to recognize and develop the bridge between biosystem of living organisms and physical systems condense matter particles. In this manner, we use the inspiration from nature for better understanding the particles physics motions, which could be observed as biomimetic system.

Every fractal object (FO) has its fractal dimension – Hausdorff dimension (D_H), a real number, smaller compared to the geometric dimension of the FO minimal space. Influence of pores' and grains' fractal dimension on materials properties has already been established [20-25]. Unlike ideal, natural fractals have scale-dependent fractal dimension [26]. Nature objects, such as water surface, air particles, trees and many others do not show intrinsic structural, but statistical self-similarity; therefore, they are considered as almost fractals. Reconstruction of such shapes requires modified mathematical approach, offered by fractal geometrical analysis. This mathematical technique can be performed on field emission scanning electron microscopy (FESEM) images, by identifying fiber phase and pores shapes and boundaries, as well as fiber-matrix bonding at the interface. In this study, fiberglass mat was used for the reinforcement of epoxy. FESEM image of enlarged fiber after the composite fracture was used for the reconstruction of data.

II. The Method

A. Fractal reconstruction of data

Fractal nature analysis of experimentally determined physical properties is performed using a novel affine fractal regression model described by the equations published in our previous research. The aim is to find coefficients that fit experimental data for the following equation system:

$$\varphi\left(\frac{x+j}{p}\right)=a_j\varphi(x)+b_jx+c_j \tag{1}$$

where $x\in[0,1)$, $0\leq j\leq p-1$, a_j represent fractal and b_j directional coefficients, with $0<|a_j|<1$, with domain [0,1), p stands for fractal period. Real solution equation system is called fractal function $\varphi:[0,1)\rightarrow R$, having mathematical fractal structure – function graph plot represents fractal curve. Higher a_j appear in the case of strong fractal oscillations. The curve fractal level defined by the equation system is L; the first fractal level is replicated in the entire domain over every of the p sub-intervals, building the second fractal level.

In order to obtain coefficients that fit the data, explicit solution of the problem that depends on the p-expansion of numbers in [0,1) is used. For L=2, this solution is

$$\varphi(0)=\frac{c_0}{1-a_0} \tag{2}$$

$$\varphi\left(\frac{\xi_1}{p}\right)=a\xi_1\frac{c_0}{1-a_0}+c\xi_1,\xi_1\neq 0 \tag{3}$$

$$\varphi\left(\frac{\xi_1}{p}+\frac{\xi_2}{p^2}\right)=a\xi_1\left(a\xi_2\frac{c_0}{1-a_0}+c\xi_2\right)+ b\xi_1\frac{\xi_2}{p}+c\xi_1,\xi_2\neq 0 \tag{4}$$

For obtaining the best coefficients, the theoretical approach computes the SSR - sum of square residuals in between the formal definition and the real values. Afterwards, the partial derivatives of SSR are equalled to zero, for minimizing the error. The best solution of the problem is given when:

$$\frac{\partial SSR}{\partial a_j}=0,\frac{\partial SSR}{\partial b_j}=0,\frac{\partial SSR}{\partial c_j}=0 \tag{5}$$

for all j=0,1,2,…,p-1. This is a problem with 3p parameters, to estimate where the equations to solve are nonlinear.

The mathematical analytical solution of this partial derivative system is not possible to compute, and a numerical approach is needed. With the software for numerical computation of the solution, called Fractal Real Finder, we worked on samples and obtained estimated curves and estimates of Hausdorff dimension. With the input of the real data, the program executes simulations and gives an output with a fractal curve as modelled above.

With the estimated fractal curves, we may estimate the Hausdorff dimension. The Hausdorff dimension is an indicator of the chaotic/irregular data behavior. The classical dimension is represented by integer: 1 for lines and curves, 2 for 2D objects, 3 for solid 3D objects. There are structures that have characteristics in between two integer dimensions. In that case, we may estimate a non-integer dimension. The fundamental theoretical mathematical non-integer dimension is Hausdorff dimension, sometimes referred as fractal dimension. The box dimension is a simplified indicator that provides estimates for the real Hausdorff dimension of real data.

Proposition. The Hausdorff dimension D of the function graph, φ solution of the above system is upper bounded by the solution of:

$$\sum_{j=0}^{p-1}\beta_j^{\,D}=1 \tag{6}$$

$$where\ \beta_j=max\left\{\frac{1}{p},|a_j|\right\},0\leq p\leq p-1$$

The coefficients with fractal relevance are those a_j such that $|a_j|>1/p$.

From the following image (Figure 1), we selected a centre small part (lighter bar) and zoomed in it to a new image (Figure 2).

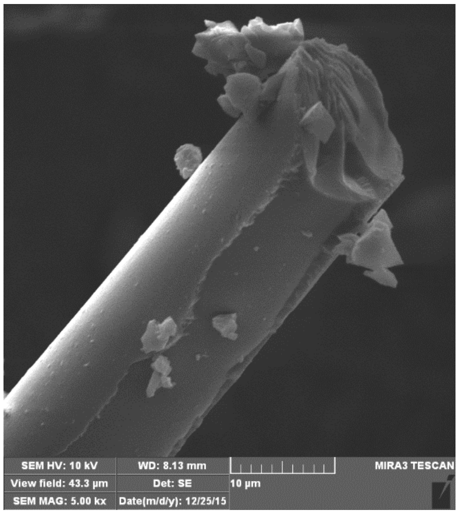


Fig. 1. FESEM of a broken glass fiber.

III. Results and Discussion

We inserted red points circling a contour of the tip of the glass fibre in a polar grid, as the following figure.

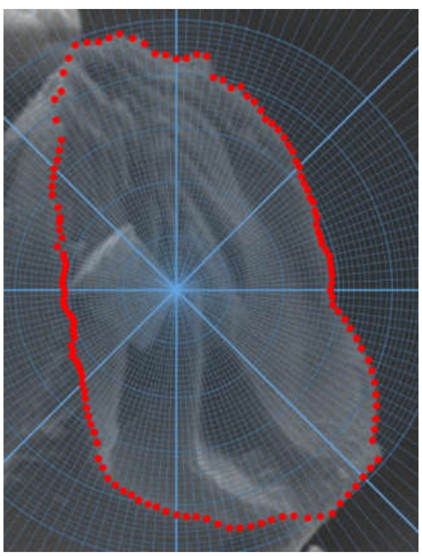


Fig. 2. Enlarged tip of the broken fiber.

We introduced red points in a border line. The fractal reconstruction is plotted in the figure next to the selected part of the image. In this case, we put the default domain in the vertical axis and upside down to match with the position of the image.

From a change of variables, from polar to cartesian variables, we run the Fractal Real Finder software for the sequence of radiuses and obtain the estimated fractal model with coefficients, as in Table below. The sequence of points is the set of radiuses corresponding to $p^L = 12^2 = 144$ angles around the image.

TABLE I
ESTIMATED COEFFICIENTS FOR THE FRACTAL CURVE OF THE RADIUS

	0	**1**	**2**	**3**	**4**	**5**
a_j	0.081	-0.044	0.012	-0.002	-0.017	0.014
b_j	0.731	-2.212	-0.805	-0.098	0.603	1.362
c_j	3.951	4.989	2.695	2.118	2.156	2.852
	6	**7**	**8**	**9**	**10**	**11**
a_j	0.027	-0.057	-0.013	0.021	-0.013	0.024
b_j	0.695	-0.005	-1.654	-0.037	0.517	0.712
c_j	3.957	5.072	4.608	2.821	2.868	3.304

This fractal reconstruction reveals no fractal coefficients (those bigger than $1/p = 0.8(3)$) and, in consequence, the corresponding Hausdorff estimate is 1. Returning to polar coordinates (radius and angle), we plot the estimated curve as follows.

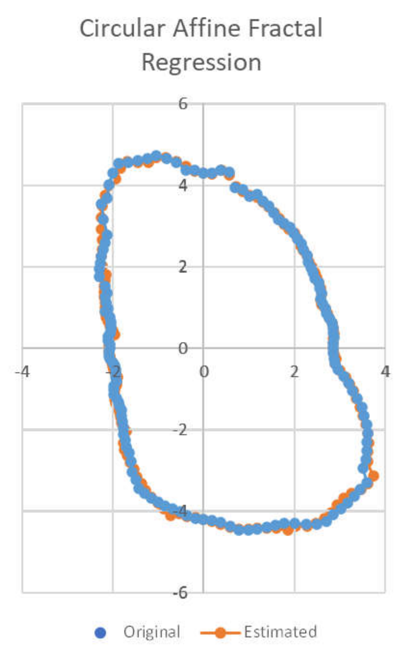


Fig. 3. Fractal curve of the fiber shape

IV. Conclusion

In this paper, fractal nature analysis was applied on fiber-reinforced composite for the reconstruction of fiber shape. The analysis with software Fractal Real Finder, fractal curve depicting the shape was obtained, as well as Hausdorff dimension of 1.21968. This indicates that the fibers have been successfully reconstructed. The finding achieved in this study enables the use of the fractal software analysis for the design and prediction of efficient reinforcement for epoxy-based composites in the future.

Acknowledgment

This work was financially supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia (Grant No. 451-03-68/2021-14/200017 and 451-03-68/2021-14/200026).

References

[1] A. K. Kaw, Mechanics of Composite Materials, CRC Press, Boca Raton, 2005.

[2] D. D. L. Chung, Composite Materials: Science and Applications, Springer-Verlag London, 2010. pp 1-34.

[3] J. N. Reddy, Mechanics of Laminated Composite Plates and Shells: Theory and Analysis,CRC Press, Boca Raton, 2003.

[4] R. Aleksić, I. Živković, P. Uskoković, Kompozitni materijali, TMF Beograd, 2015.

[5] V. Radojević, R. Aleksić, Mehanička svojstva materijala, TMF Beograd, 2015.

[6] C. Shi, Y. L. Mo, Engineering Materials for Technological Needs, vol. 1, World Scientific Publishing Co. Pte. Ltd, Singapore, 2004.

[7] I. Radovic, A. Stajcic, A. Radisavljevic, F. Veljkovic, M. Cebela, V.V. Mitic, V. Radojevic, Solvent effects on structural changes in self-healing epoxy composites. *Mater.Chem. Phys*., 256 (2020) 123761.

[8] I.M. Radovic, D.B. Stojanovic, A. Kojovic, M. Petrovic, P.S. Uskokovic, V.J. Radojevic, R.R. Aleksic, Healing efficiency of polystyrene electrospunnanofibers with Grubbs' catalyst in thermosetting composite, *J. Compos. Mater*., 51 (2017) 3003–3016.

[9] A. Grujić, J.Stajić-Trošić, M. Stijepović, S.Putić, D.Nedeljković, A.Stajčić, R.Aleksić, Dynamic Mechanical Behaviour of Polymer Bonded Nd-Fe-B Composite Materials, *Mater. Trans*., 53 (2012) 395-400.

[10] A.P. Stajčić, J.T. Stajić-Trošić, A.S. Grujić, M.Z. Stijepović, N.L. Lazić, T. Žák, R.R. Aleksić, Hybrid Nd-Fe-B/barium ferrite magnetic materials with epoxy matrix, *Hem. Ind*. 66 (2012) 301–308.

[11] A. Stajcic, I. Radovic, V. Cosovic, A. Grujic, J. Stajic-Trosic, R. Jancic-Heinemann, The Influence of Barium Ferrite Nanoparticles on Morphological and Mechanical Properties of Ethyl Cellulose Based Nanocomposites, *Sci. Sinter*., 51 (2019) 277-283.

[12] D.M. Nedeljković, A.P. Stajčić, A.S. Grujić, J.T. Stajić-Trošić, M.M. Zrilić, J.S. Stevanović, S.Z. Drmanić, The Application of Zeolite Nanopowder for the Construction of the Dense Composite Polymer Membranes for Carbon Dioxide Separation, *Dig. J. Nanomater. Bios.,* 7 (2012) 269-782.

[13] K. Chen, X. Zhao, F. Zhang, X. Wu, W. Huang, W. Liu, X. Wang, Influence of gamma irradiation on the molelucular dynamixs and mechanical properties of eoxy resin, Polym. Degrad. Stab. 168 (2019) 108940.

[14] P. Zhang and G. Li, Advances in healing-on-demand polymers and polymer composites. *ProgPolymSci* 57 (2016) 32–63.

[15] A. Zucchelli, M.L. Focarete, C. Gualandi and S. Ramakrishna, Electrospun nanofibers for enhancing structural performance of composite materials. *PolymAdvTechnol* 22 (2011) 339–349.

[16] D.D.L. Chung *Composite materials: science and application.* 2nd ed. London: Springer-Verlag, 2010, p.95.

[17] S.K. Ghosh, Self-healing Materials: Fundamentals, Design Strategies, and Applications. In: Ghosh SK (ed) *Self-healing Materials: Fundamentals, Design Strategies, and Applications*, Weinheim: WILEY-VCH Verlag GmbH & Co. KGaA, 2009, pp. 1–28.

[18] M. Scheiner, T.J. Dickens and O. Okoli, Progress towards self-healing polymers for composite structural applications. *Polymer* 83 (2016) 260–282.

[19] V.V. Mitic, G. Lazovic, V. Paunovic, N. Cvetkovic, D. Jovanovic, S. Veljkovic, B. Randjelovic, B. Vlahovic, Fractal frontiers in microelectronic ceramic materials, Ceram. Int. 45 (2019) 9679-9685.

[20] R.P. Taylor, B. Spehar, Fractal Fluency, An intimate relationship between the brain and processing of fractal stimuli, in: Di Ieva A (Ed.), The Fractal Geometry of the Brain, Springer, 2016.

[21] N.C. Kenkel, D.J. Walker, Fractals in the biological sciences, COENOSES 11 (1996) 77–100.

[22] V.P. Dimri, Application of Fractals in Earth Sciences, Balkema, Rotterdam, 2000.

[23] L.G. Lanza, J. Gallant, Fractals and Similarity Approaches in Hydrology, Part 1. Theory, Organization and Scale, Published Online John Wiley and Sons, 2006.

[24] N.Cohen, Fractal antennas and fractal resonators, Patent US 6452553 B1, 1995.

[25] C.P.Baliarda. et al.,Space-filling miniature antennas, PatentUS 8207893 B2, 2000.

[26] V.V. Mitic, Lj. Kostic, V. Paunovic, G. Lazovic, M. Miljkovic, Fractal nature structure reconstruction method in designing microstructure properties, Mater. Res. Bull. 101 (2018) 175-193.

The Neural Network Application on Ceramics Materials Density

Srdjan Ribar, Vojislav V.Mitic, Branislav Randjelovic, Dusan Milosevic, Vesna Paunovic, Hans Jorg Fecht, Branislav Vlahovic

***Abstract* — In this research back propagation neural network (BP) was applied on ceramics material samples, consolidated by sintering data obtained in analyzed experiment in a specific way. The main characteristic of BP is that it is capable to perform arbitrarily input-output data mapping due to large set of adjustable coefficients called weights. Desired mapping is possible to achieve if coefficients are set to proper value and this procedure is called training. At the beginning of training process weights are set to random values. Error is defined as a difference between desired and actual network output and weight coefficients have a contribution in generating error.**

Within experimental from material density values sintering results, measured on a surface, we investigate a possibility to calculate density within sintered structure. In this case BP training procedure is used as a tool to spread values measured on a sample surface – density. In this investigation network errors are replaced with density values obtained in ceramics sintering process. We sucesfully performed this neural network application novelty in ceramics material sciences within sintering process for the case $\rho=5.4x10^3[kg/m^3]$.

***Index Terms* — ceramics, sintering, neural network, error, density.**

I. Introduction

In this research ceramics material samples, consolidated by sintering data obtained in analyzed experiment are explored using back propagation neural network. Back propagation neural network is performing arbitrarily input-output data mapping due to large set of adjustable coefficients, Those coefficients are called weights. If coefficients are set to proper value, then desired mapping can be achieved and this procedure is called training of neural network. At the beginning of this process weights are set to random values, and during training process input-output training data is set to a network. It means that desired input-output mapping is known.

Due to inappropriate weight values mapping is performed with significant error, where error is a difference between desired and actual network output. All weight coefficients have a significant contribution in generating this error. Since this measured error is occurred due to incorrect weight values of a whole network it is necessary to change weights through the whole network in the sense to minimize it. This process starts from the output layer where error is measured and finishes in input layer. This procedure of changing weight values is called training. Training is applied on a whole training set data nummerous times. With the change of the values of the coefficients, through the training, process error decreases and after trainining process is finished further on network mapping is satisfactory even for a new input data on network has not been trained. Training process implies changing all network coefficients starting from network output through the whole network to the input. Since this process performs from network output to network input it is called error back propagataion [1,2,3,4]. So, contribution of all network elements generating output error is calculated. Training process is finished when all input-output data are mapped within predefined error.

The main idea in this paper is to extend the neural network application on processes in sintering and calcualtion of various parameters, within different sintering temperatures intervals. The process of consolidation the ceramics materials at different thermal conditions has a very important relation to density.

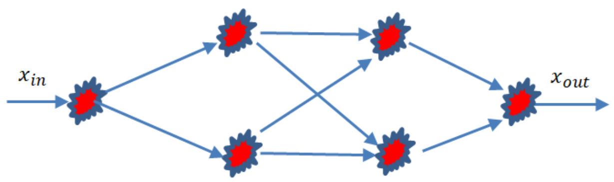


Figure1. The principle scheme neural networks.

Srdjan Ribar is with Faculty of Mechanical Engineering, University of Belgrade, Kraljice Marije 16, 11120 Belgrade Serbia, Belgrade, Serbia (e-mail: srdjanribar@gmail.com).

Vojislav V.Mitić is with Faculty of Electrical Engineering, University of Nis, 14 Aleksandra Medvedeva, 18000 Belgrade, Serbia (e-mail: vmitic.d2480@gmail.com).

Branislav M.Randjelović is with Faculty of Electrical Engineering, University of Nis, 14 Aleksandra Medvedeva, 18000 Belgrade, Serbia (e-mail: bane@elfak.ni.ac.rs).

Dusan Milosević is with Faculty of Electrical Engineering, University of Nis, 14 Aleksandra Medvedeva, 18000 Belgrade, Serbia (e-mail: dusan.milosevic@elfak.ni.ac.rs).

Vesna Paunović is with Faculty of Electrical Engineering, University of Nis, 14 Aleksandra Medvedeva, 18000 Belgrade, Serbia (e-mail: vesna.paunovic@elfak.ni.ac.rs).

Hans J. Fecht is with University of Ulm, Institute of Functional Nanosystems, Ulm, Germany (e-mail: hans.fecht@uni-ulm.de).

Branislav Vlahovic is with North Carolina Central University, (NCCU), Durham, North Carolina USA (e-mail: vlahovic@nccu.edu).

We introduce neural networks (Fig.1) as an efficient tool for application on calculation of different physical parameters. They are very useful, from experimental point of view, because their applciation on results of measurements fit much more extended experimental intervals with exeptible neglected error.

Material structure is assumed as a multi thin layers coating around the both sides grains interconnections. Any signal measured on the material surface could be propagated instead neural network error through the whole structrure. This idea was analyized in [5,6,7] where relative capacitance measured on a sample surface was propagated through the ceramics structure assuming that ceramic structure can be presented by a neural network, Fig. 1.

Sintering material density value (in last column of Table I) is one of experimentally obtained parameters. These values are measured on a surface. Our goal is to investigate a possibility to calculate density of a sintered material within a sintered structure. In order to solve this problem a possible solution would be a usage of neural networks. Back propagation neural network (BP) seems as an appropriate tool taking into account a training procedure that changes weight coefficient values [8,9,10,11]. Weight values are changed all through whole network proportionately to an error value measured on an output network layer.

In this case BP training procedure is used as a tool to spread values measured on a sample surface – density. Various BP structures are trained to map arbitrary input-output data. During training process an absolute error value measured on an output is used to change weight values through the network. In this investigation network errors are replaced with density values obtained in sintering process, Table 1. first row $\rho=5.4x10^3[kg/m^3]$.

II. Experimental results and methods

The ceramic powder preparation for sintering consolidation for $BaTiO_3$ - ceramics samples is consisted of different steps: (a) measuring and forming a mixture of starting powders with impurities, (b) wet mixing and spraying, (c) molding and process control and (d) preparation, samples sintering and process control. We applied high purity commercial $BaTiO_3$ Murata powder [12] (mean grain size <2μm, 99.9 % purity). We analyzed the influence of sintering parameters on the final $BaTiO_3$-ceramics characteristics, onto the density. Powder mixture was processed into a mill with balls and water and organic binders were added and homogenization was about 48 hours, and the mass was transferred by a membrane pump and dried, so we got desired powder granulation. We tested the material density every hour by a special vessel and after that vibrating sieve was applied. Roughly shaped powder particles were of diameters 10-130μm.

We analized the sintering temperatures (1190-1370°C) and time (2-3h) with impact of additive CeO_2, $MnCO_3$. For this research paper we stress the relation for pressure 86MPa and density.

TABLE I
EXPERIMENTAL RESULTS

sample type	***P [MPa]***	***ρ [kg/m³]***
$BaTiO_3$ – ceramics with basic mixture	86	*5.4x10³*
$BaTiO_3$ -ceramics: composition 0.1%CeO_2+0.14%$MnCO_3$	86	*3.2x10*
$BaTiO_3$ -ceramics: composition 0.1%CeO_2+0.14%$MnCO_3$	86	*3.4x10*

In further theoretical experiment we used concrete data $\rho=5.4x10^3[kg/m^3]$, from the first raw of the Table I.

A. Theoretical experiment and neural network method application

For a network with one neuron in each of two hidden layers errors calculated (Table II) in training process are

TABLE II

neuron	first hidden layer	second hidden layer	output neuron
1	*0.076786*	*0.141702*	*0.185405*

Calculated density ρ in hidden layers (Table III) is

TABLE III

neuron	first hidden layer	second hidden layer	output neuron
1	*2200*	*4300*	*5400*

Error for a network with two neurons in a first hidden layer and one neuron in a second hidden layer (Table IV) is

TABLE IV

neuron	first hidden layer	second hidden layer	output neuron
1	*0.023776*	*0.314802*	*0.887293*
2	*0.045009*		

Calculated density ρ in hidden layers (Table V) is

TABLE V

neuron	first hidden layer	second hidden layer	output neuron
1	*273,9*	*1900*	*5400*
2	*144,6*		

Error for a network with two neurons in a first hidden layer and two neurons in a second hidden layer (Table VI) is

TABLE VI

neuron	first hidden layer	second hidden layer	output neuron
1	*-0.0654*	*-0.2297*	*0.903708*
2	*-0.05848*	*-0.05259*	

Calculated density ρ in hidden layers (Table VII) is

TABLE VII

neuron	first hidden layer	second hidden layer	output neuron
1	*349*	*314*	*5400*
2	*391*	*1372*	

Error for a network with three neurons in a first hidden layer and one neurons in a second hidden layer (Table VIII) is

TABLE VIII

neuron	first hidden layer	second hidden layer	output neuron
1	*-0.04998*	*-0.10846*	*0.668582*
2	*-0.06411*	*-0.16845*	
3	*-0.03064*		

Calculated density ρ in hidden layers (Table IX) is

TABLE IX

neuron	first hidden layer	second hidden layer	output neuron
1	247	*1360*	*5400*
2	518	*876*	
3	*404*		

There is a possibility to calculcarte error and ednsity using other neural networks (for example: 3 neurons in first level, 2 neurons in second level; 3 neurons in first level, 3 neurons is second level; 4 neurons in first level, 1 neuron is second level) but this will be part of some future researche.

B. Results and discussion

Obtained and trained neural networks are given on next figures.

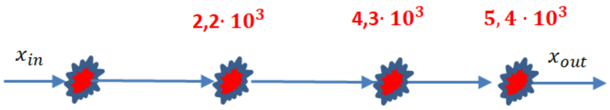


Figure 2. Neural network with one neuron per each hidden layer.

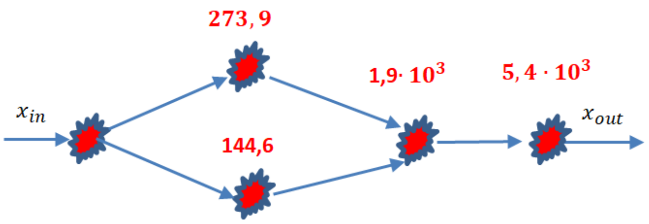


Figure 3. Neural network with two neurons in first hidden layer and one neuron in second hidden layer

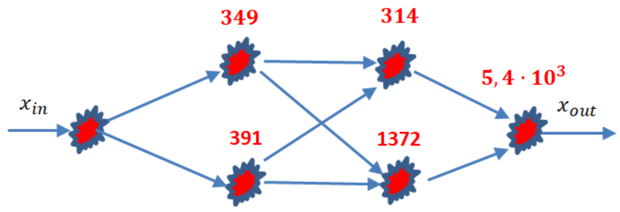


Figure 4. Neural network with two neurons per each hidden layer.

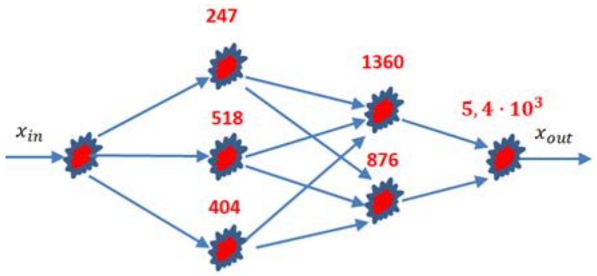


Figure 5. Neural network with three neurons in first hidden layer and two neurons in second hidden layer.

Based on experimentaly consolidated samples and applied the neural networks in three diferente cases shown on Fig.2, Fig.3 and Fig.4, we sucessfully performed this original novelty in getting the samples surface density, based on a theoretical experiment and neural networks calculations.

The advantage in this methodology is within the collecting the ceramics materials densities, from the sample surfaces, based on neural networks. This is efficient and productive way for data densities which is avoiding much more experimental activities and materials and energy and time losses.

III. CONCLUSION

In this research report we explained the neural network method and its application on some results, which are surface density on the experiment. We collected in our experiment a plenty of different results, but we consistently demonstrated this novelty with surface density $\rho = 5.4x10^3 [kg/m^3]$..

We presented the densities values within the different ceramics sample microstructure levels. Without this method, we could have problem to calculate the desired material density information only by stochastic mathematical approach. On this way, we have very new frontiers in science of sintering ceramics processing and technology howe can get the densities within the whole morphology.

All of this also opens new directions for predicting - designing and prognosis within the ceramics structure, where we would like to form the projective structures on a quit precise way [13,14,15,16,19].

In further experiments and researche we will show how those results, based on surface density, can be calculated on some different neural networks [16,17,18,20], with different shapes and various combination of neurons in hidden layers. We will also show how this approach can be applied for calculation of various other parameters in similar manner, based on obtained experimental results.

REFERENCES

[1] Rumelhart, D., *et al.* Learning representations by back-propagation errors, *Nature*, 323, (1986) pp. 533-536

[2] da Silva I.N., et al. Artificial Neural Network Architectures and Training Processes, in: *Artificial Neural Networks*. Springer, Cham., 2017, pp.21-28

[3] Ribar S. *et al*. Neural Networks Application on Human Skin Biophysical Impedance Characterisations, *Biophysical Reviews & Letters*, 16, (2021), pp. 2021 – 19

[4] Hecht-Nielsen: Theory of the backpropagation neural network, International Joint Conference on Neural Networks, IJCNN, Washington, D. C., USA, 1989. Vol. 1, pp 593-605

[5] Mitic V. V., *et al*. The Artificial Neural Networks Applied for Microelectronics Intergranular Relations Determination, *Integrated Ferroelectics*, 212, (2020)

[6] Mitic V. V., *et al*. Neural Networks and Microelectronics Parameters Distribution Measurements Depending on Sintering Temperature and Applied Voltage, MPL B, 34 (**35)** (2020) pp. 2150172.

[7] Randjelovic B., *et al*. “Ceramics, Materials, Microelectronics and Graph Theory new Frontiers”, MPLB, Vol 34 (34) (2020) pp.2150159.

[8] Mitic V. V., *et al*. Entropy and fractal nature, *Thermal Science*, 24 (2020) 3B pp. 2203–2212.

[9] Mitic V. V., *et al*., “Butler-Volmer Current Equation and Fractal Nature Correction in Electrochemical Energy”, *Thermal Science***,** (2020), p. 232-232,

[10] Randjelovic B., *et al*. “Graph Theory Approach in Synthetized Diamonds Electrophysical Parameters Defining”, in: *Bioceramics, Biomimetic and other Compatible Materials Features for Medical Applications,* (Ed. Najman S., Mitic V., Groth T., Barbeck M., Yu Chen P., Sun Z.), Springer Nature, Cham, Switzerland, 2021, accepted for publication

[11] Ribar S., *et al*. “Neural Networks from Biophysical Applications in Microelectronics Parameters Measurements”, in: *Bioceramics, Biomimetic and other Compatible Materials Features for Medical Applications,* (Ed. Najman S., Mitic V., Groth T., Barbeck M., Yu Chen P., Sun Z.), Springer Nature, Cham, Switzerland, 2021, accepted for publication

[12] Mitic V.V.: Structure and electrical properties of $BaTiO_3$ ceramics, (in Serbian), Zaduzbina Andrejevic, Belgrade, Serbia, 2001

[13] V.V.Mitic, B.Randjelovic, I.Ilic, S.Ribar, **C**.An-Lu, A.Stajcic, B.Vlahovic.“The 3D graph approach for breakdown voltage calculation in BaTiO3 ceramics”, IJMPB, Vol 35 (7) (2021) pp. 2150103.

[14] V.V.Mitic, G.Lazovic, B.Randjelovic, V.Paunovic, I.Radovic, A.Stajcic, B.Vlahovic: “Graph Theory Applied to Microelectronic Intergranular Relations”, Ferroelectrics, 570 (2021) p. 145-152.

[15] V. Mitlić, G. Lazović, B.M. Randjelović, V. Paunović, J.M. Wu D.Mancić,, J.R. Hwu, “Fractal Microeletronic Frontiers and Graph Theory Applications”, Int.Conference MS&T **2019**, Book of Abstracts, Portland, USA, 29 September – 03 October 2019.

[16] V. Mitić, G. Lazović, C.A.Lu, I.Radovic, V. Paunović, A. Stajcic, B.Randjelović, S.Ribar, B. Vlahovic, “Investigation of intergranular dielectric properties within the relation between fractal, graph and neural networks theories”, Int.Conf. Electronic Materials and Applications EMA-2021, Virtual, January 19-21, 2021.

[17] V. Mitić, C. Serpa, C.A.Lu, V. Paunović, I.Radovic, B.Randjelović, S.Ribar, A. Stajcic, B. Vlahovic, “Fractal, graph and neural network theories applied on BaTiO3 electronic ceramics”, Int.Conf. Electronic Materials and Applications EMA-2021, Virtual, January 19-21, 2021.

[18] V. Mitić, I.Radovic, C.A.Lu, S.Ribar, B.Randjelović, A. Stajcic, H. Fecht, B. Vlahovic, “Neural networks and applied graph theory approaches for intergranular properties measurements investigation”, Int.Conf. Electronic Materials and Applications EMA-2021, Virtual, January 19-21, 2021.

[19] V. Mitić, C.A.Lu, B.Randjelović, V. Paunović, I.Radovic, A. Stajcic, J.Kotnik, H.Fecht, B. Vlahovic, “3D graph theory application on modified nano BaTiO3 electronic ceramics”, Int.Conf. Electronic Materials and Applications EMA-2021, Virtual, January 19-21, 2021.

[20] Mitic V. V *et al.* Sinterign Temperature Influenceo on Grains Function Distribution by Neural Network Application, Thermal Science (2021) (accepted).

Structural Characterization of La($Mg_{1/2}Ti_{1/2}$)O_3 (LMT) Perovskite for Mobile communications

Kouros Khamoushi, Vojislav Mitić, Jelena Manojlović, Vesna Paunović, *Member, IEEE*, Goran Lazović

***Abstract*—The phase group and structure properties of La($Mg_{1/2}Ti_{1/2}$)O_3 (LMT) ceramics, prepared via the mixed oxide route, were examined in this research. A single-phase La($Mg_{1/2}Ti_{1/2}$)O_3 ceramic were produced at different sintering temperature from 1250ºC to 1675ºC. The heating rate was 25 ° C to the sintering temperature and the cooling rate from 2 ° C per minute to room temperature. The XRD analysis determined that LMT ceramics have a cubic crystal structure with a lattice parameter a = 0.392 nm. The theoretical density of this ceramics is 6.0846 g/cm3. These materials must be sintered from 1550-1675°C to achieve a sintering density of about 99% of theoretical density. At a temperature lower than 1500°C a density was 93% of theoretical density. The temperature coefficient of resonant frequency for LMT was -72 ppm/ºC, and the Quality factor was 34000 at a frequency of 8.07 GHz.**

Key words- Microstructure: Grain growth; single phase La($Mg_{1/2}Ti_{1/2}$)O_3; Lanthanum magnesium titanium oxide.

1. Introduction

Ceramic components made of dielectric material are important for the operation of filters and oscillators in several microwave systems, such as military radar systems, mobile communications, and satellite TV receivers. Perovskites have a cubic structure with the general formula ABO3, and many are based on BaTiO3. If barium ions are located at the corner of the cube, oxygen is located at face-centered sites, and titanium ions occupy the body-centered sites. Distortion of this sort of compound produces an electrical signal, permitting BaTiO3 to serve as a transducer [1]. The size of the ions has an important share in the stability of the crystal structure, and low distortions can result in lower symmetry and drastically change the property of the compound [2]. An important class of perovskites is multiferroic, which simultaneously shows ferromagnetic, ferroelectric, and ferroelastic parameters in the same segment. Another advantage of perovskite compound is in medical imaging which they can be ten times more sensitive when is used for X-ray detector [3].

In microwave communications, a dielectric resonator filter is used to discriminate between wanted and unwanted signal frequencies in the transmitted and received signal. When the wanted frequency is extracted and detected, it is necessary to maintain a strong signal nevertheless. For clarity, it is also critical that seasonal temperature changes do not affect the expected signal frequencies, and resonator materials for practical application should have reliable properties. A high relative dielectric constant is needed so that the materials can be miniaturized, and a high-quality factor (Q) for improved selectivity. Low-temperature variation of the material's resonant frequency is also required so that the microwave circuits remain stable. Everything from the electromagnetic properties to the microstructure of the material is important for the result.

The gain to reduce the size, load, and expense of microelectronic tools rules for tuning τ_f in multipart perovskites have already been created, accompanied by the piezoelectric materials the most leading material is Pb(Zr_x Ti_{1-x}) or PZT which have much application in ultrasound transducer ceramics which used also as capacitor [4,5] according to the work [6] that τ_ε in Ba- and Sr-based complex perovskites are profoundly connected to the start and degree of octahedral tilting. Moreover, it can be tuned through ±300 MK^{-1} without significantly adjusting Q or dielectric constant (ε_r) by manipulating the perovskite tolerance factor, *t*,

$$t = \frac{R_A + R_O}{\sqrt{2}\,(R_B + R_O)} \quad (1)$$

from 1.01 - 0.93, where R_A, R_B, and R_O are the radii of the ions in the perovskite (ABO_3) structure. Reducing *t* results in the initiation of octahedral tilt modifications. The link between τ_ε and τ_f is:

$$\tau_f = -\left(\tau_\varepsilon / 2 + \alpha_L \right) \quad (2)$$

Kouros Khamoushi - Natural Science, Tampere University, P.O. Box 1001 FI-33014Tampere, 33720/ Finland, (Email:kouros.khamoushi@tuni.fi).

Vojislav Mitić - University of Niš, Faculty of Electronic Engineering,, Aleksandera Medvedova 14 Niš 18000 Serbia, (Email: vmitic.d2480@gmail.com).

Jelena Manojlović- University of Niš, Faculty of Mechanical Engineering,, Aleksandera Medvedova 14 Niš 18000 Serbia, (Email: jmanojlo@gmail.com)

Vesna Paunović – University of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Srbija (email: vesna.paunovic@elfak.ni.ac.rs).

Goran Lazović- University of Belgrade, Faculty of Mechanical Engineering, Serbia, (email: goran.lazovic@gmail.com)

where α_L is the coefficient of linear thermal expansion (≈10 MK^{-1} for perovskites).

$La(Mg_{1/2}Ti_{1/2})O_3$ is a perovskite formed in the pseudocubic system, although the exact crystal structure has not been fully elucidated. Its tolerance factor is t = 0.95, which, approves to the work [7], reveals the existence of both in-phase and anti-phase tilting of oxygen octahedra. The impacts of this tilting have been detected by XRD in the form of ½(111) ordering at 1600°C sample.

Superlattice reflections corresponding to anti-phase and in-phase tilting, respectively. Systematic cation displacement was also detected by the presence of ½(210) and ½(111) reflections were observed and attributed to the expansion of the unit cell caused by the B-site cation changes. It is possible by the present tilts to reveal two possible orthorhombic space groups (Pbmn or $P2_1/n$) for the structure, but Walsenburg photographs or TEM work would be required to establish the structure.

Structurally, tilting of octahedra has similar effects as cation ordering in 1:1 type complex perovskites. Both result in expanding unit cells [8] even though cation order also reduces the space available in which the A-site cation types can rattle, it depends on the radii size of rear earth elements used in the compound. In this study, the structural characteristic of $La(Mg_{1/2}Ti_{1/2})O_3$ (LMT) in various sintering temperatures for the application of mobile phones and telecommunications was investigated.

2. Experimental Procedure

In this work, conventional mixed oxide powder processing techniques were used for the preparation of samples. A detailed description of the laboratory procedure for sample preparation is given in references [9] and [10]. Starting materials included La_2O_3 (99.9% Meldform Rare Earths, U.K), TiO (99.8% Alfa Aesar, U.K.), and $(MgCO_3)_4Mg(OH)_25H_2O$ (99% Aldrich Chemical Company, Inc, USA). La_2O_3, the rare-earth oxide, was first purposely hydrated in distilled water to form $La(OH)_3$. These hydrates were then used in the subsequent processing method, which involved milling stoichiometric quantities of powders together in a porcelain mill pot partly filled with ZrO_2 media and distilled water for four hours.

A small amount (1wt%) of Dispex A40 (Allied Colloids, Bradford, U.K.) was added as a deflocculant. The slurries were then dried overnight at 80°C. Dried powders were subsequently granulated with a mortar and pestle and sieved to under 250 μm. Calcination was achieved using a two-stage process. First, the powder was heated to 650 ºC for 2 hours in an open Al_2O_3 crucible to ensure dihydroxylation of the $La(OH)_3$. The completion of the dihydroxylation reaction was monitored by measuring the weight loss at this stage of the process. Second, this same powder was combined by hand, a lid was placed over the crucible, and it was re-heated to 1400ºC for 2 hours. Afterward, the powder was re-milled for a further four hours with 2wt% PEG 1500 (Whyte Chemicals, London) being added in an aqueous solution 5-15 min before completion. These slurries were then dried and granulated as above and subsequently pressed (125 MPa) into cylindrical pellets 10 mm in diameter and 3 mm thick. Sintering was conducted in closed alumina boats for 6 hours at temperatures ranging from 1125°C to 1675°C. Pellets were weighed up before and after sintering to quantify the degree of deficiency. Dense and homogeneous ceramic powder of lanthanum magnesium titanium oxide was obtained after pressing powder 125 MPa and sintering at 1600°C for 6 hours.

Phase groupings were verified by scanning electron microscopy (JSM 6300, Jeol, Tokyo) and x-ray diffraction (D50000, Siemens, Germany), using CuKa radiation. Some samples underwent thinning by ion milling (model 600, Gatan, California, USA) for observation in the transmission electron microscope (JEM 2010, Jeol, Tokyo).

3. Results and Discussion

Figure 1 shows the change in density of La $(Mg_{1/2}Ti_{1/2})O_3$ ceramics as a function of sintering temperature. The density of LMT ceramics increased with increasing sintering temperature and reached (87% TD) at 1250°C. Until the sintering temperature of 1300°C, the densities do not change, and then increases to (98% TD) at 1500°C. At this temperature, high-quality, high-density material is obtained.

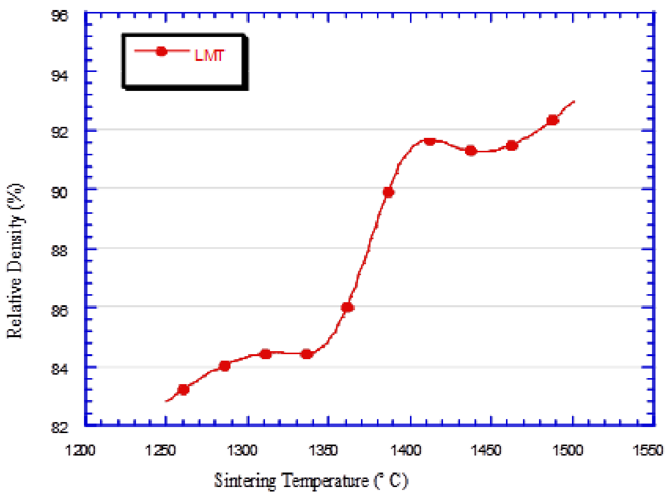


Fig. 1. Relative density of LMT samples in function of sintering temperature (Tsin=1250-1500 ºC).

XRD patterns of LMT samples, sintered at 1400-1675ºC temperature, were shown in Fig 2. Eleven samples were selected, and starting sample was sintered at 1400 ºC for 6 hours and this sintering process continued for each individual sample by increasing 25 ºC to the previous temperature. This successive test continued for every sample and ended at 1675ºC. All the samples show perovskite crystal structure. Peaks from 1400-1675ºC match the LMT structure. All the peaks match PDF card numbers 49-242. Crystallography search match indicates that crystal structures of samples are cubic with a space group of Pa3.

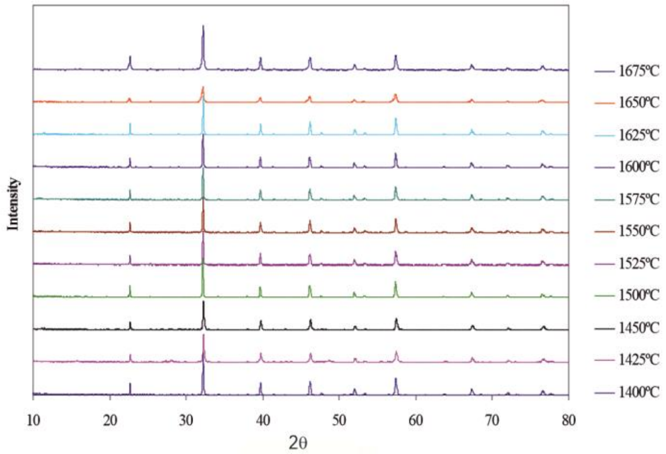


Fig. 2. X-ray diffraction (XRD) of LMT for sintering temperature of 1400-1675ºC.

The SEM image of $La(Mg_{1/2}Ti_{1/2})O_3$ is shown in Figure 3. The image shows an LMT sample sintered at 1650°C. As can be seen from Figure 3, the sample is a single-phase high-dense ceramics. The material with the second phase must show a different contrast. The background scattered image of these samples showed no differences. If the samples show two different colors in contrast, then the investigated material is not single-phase and has impurities. However, several samples of LMT were tested, and all results show pure and very dense LMT material. Also, these samples are characterized by polygonal grains the size ranging from 1 to 6µm.

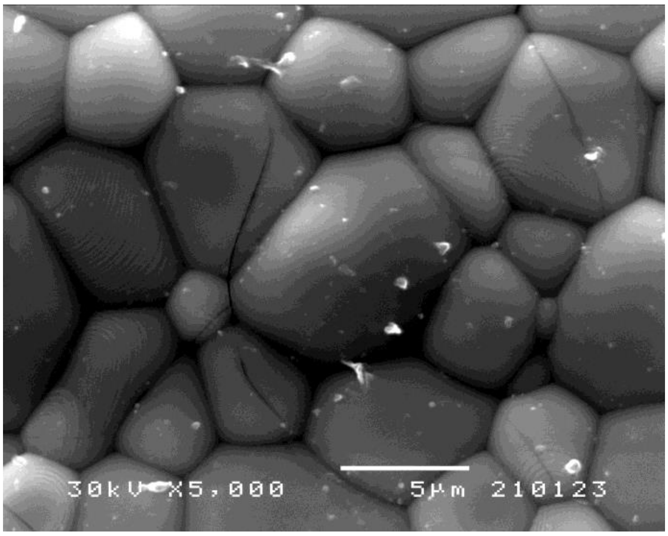


Fig. 3. SEM microphotography of LMT ceramics, sintered at 1650 °C, beam direction [111].

3.1. STRUCTURE OF $La(Mg_{1/2}Ti_{1/2})O_3$

The X-ray pattern for LMT powder shown in figure 4. It is indexed according to the Magnesium Lanthanum Titanium Oxide PDF card number 49-242 with single phase cubic crystal structure with $a \approx 0.3.92$ nm. XRD of sample powder did not show any broadening or splitting of cubic peaks. All the XRD taken from sample of calcined as well as the sintered pellets were similar.

Nevertheless, to avoid the uncertainty for the definition of the exact crystal structure and space group of LMT samples, high resolution transmission electron microscopy and Rietveld refinement by using GSAS should be performed, this work is under consideration for future.

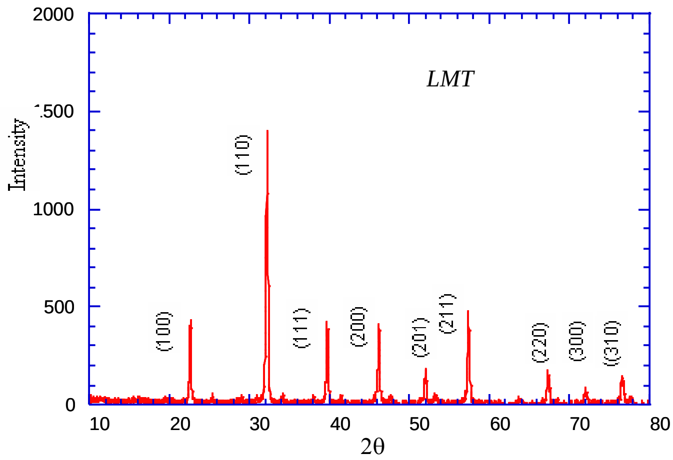


Fig. 4. X-ray diffraction pattern of $La(Mg_{1/2}Ti_{1/2})O_3$ powder after sintered at 1600°C. All the peaks have been indexed according to PDF card numbers 49-242

The temperature coefficient of the microwave resonant frequency (τ_f) and the quality factor (Q) at the resonant frequency were made from the dielectric characteristics.

The results of $Q \times f$ measurements illustrate that LMT ceramics have quality factors $Q \times f = 34{,}000$ at frequency 8.07 GHz. The temperature coefficient of the resonant frequency of LMT is $\tau_f = -72$ ppm/ ºC.

4. Conclusions and Future Work

In this paper the structure characteristics of $La(Mg_{1/2}Ti_{1/2})O_3$ were explored. LMT has lattice parameter a = 0.392 nm and it shows a cubic crystal structure. The theoretical density is 6.0846 g/cm^3. This material can be sintered from 1550-1675 ºC to achieve sintering density of about 99% (1550ºC and 1600ºC). At lower temperature the density was under 97% that of theoretical density. The temperature coefficient of resonant frequency for LMT was -72 ppm/ ºC. The quality factor for LMT was 34000 where saturated at frequency 8.07 GHz. Clearly these materials show a good potential as filters for mobile microwave telecommunications.

Acknowledgements:

This work has been supported by Queen Mary University of London, UK, Tampere University of Technology, Finland, and the Ministry of Education, Science and Technological Development of the Republic of Serbia (Ev. No. 451-03-9 / 2021-14 / 200102).

REFERENCES

[1] K. A. Muller and T. W. Kool, properties of Perovskite and other oxide, world Scientific Publishing, Singapore (2010).

[2] H. Tanaka, H. Tabata, K. Ota, T. Kawai, "Molecular dynamics prediction of structural anomalies in ferroelectric and dielectric $BaTiO_3$-$SrTiO_3$-$CaTiO_3$ solid solutions" Physical ReviewVol.53, no 21, p. 14112-14113, 6th Ed. (1996).

[3] YC Kim, KH Kim, DY Son, DN Jeong, JY Seo, YS Choi "Printable organometallic perovskite enables large-area, low-dose X-ray imaging,", Nature, 550, 87 (2017).

[4] C. Rossel1, M. Sousa1, S. Abel1, D. Caimi, A. Suhm , J. Abergel,G. Le Rhun, and E. Defay. Defay. "Temperature dependence of the transverse piezoelectric coefficient of thin films and aging effects" Journal of Applied Physics, 115, 034105 (2014).

[5] Yingfen Wei, Pavan Nukala, Mart Salverda, Sylvia Matzen, Hong Jian Zhao, Jamo Momand,Arnoud S. Everhardt, Guillaume Agnus, Graeme R. Blake, Philippe Lecoeur, Bart J. Kooi, Jorge Íñiguez, Brahim Dkhil, "A rhombohedral ferroelectric phase in epitaxially strained Hf0.5 Zr0.5O2 thin film "Journal of Nature materials (2018) 17,pages 1095–1100 (2018).

[6] Kouros. Khamoushi, Vojislav. V. Mitic "Comparison between crystal structire and dielectric properties $Nd(Mg_{1/2}Ti_{1/2})O_3$ (NMT) and $Nd(Zn_{1/2}Ti_{1/2})O_3$ (NZT)," In process accepted on 11.5.2021 Journal of Modern Physics Letters B, (2021).

[7] D.Y.Lee S.J.Yoon, J.H.Yeo, S.Nahm,J.H.Paik, K.C.Whang, and B.G.Ahn,"Crystal structure and Microwave Dielectric properties of La(Mg½ Ti½)O3 ceramics," Journal of Materials Science Letter, vol.19 pp131-134(2000).

[8] M.P Seabra, A N Salak, M Avdeev and M Ferreira., "Structure and dielectric characterization of the $La(Mg_{½}Ti_{½})O_3$-$Nd(Mg_{½}Ti_{½})O_3$ system," Journal of Physics Condensed Matter,", 15, (2003).

[9] R. Ubic, K. Khamoushi, Hu, and Isaac Abrahams, "Structure and properties of La(Mg0.5 Ti0.5)O3," Journal of European Ceramic Society, No.26, 1787-1790 (2005).

[10] R. Ubic, K. Khamoushi, D. Iddles, T. Price, "Processing and dielectric properties of $La(Mg_{0.5}Ti_{0.5})O_3$ and Nd $Mg_{0.5}Ti_{0.5})O_3$," Ceramic transaction, 167, 21-3, (2004)

НУКЛЕАРНА ТЕХНИКА
/
NUCLEAR ENGINEERING AND TECHNOLOGY

(НТ/NTI)

Posledica merenja brzih napona Kerovim efektom u polju gama zračenja

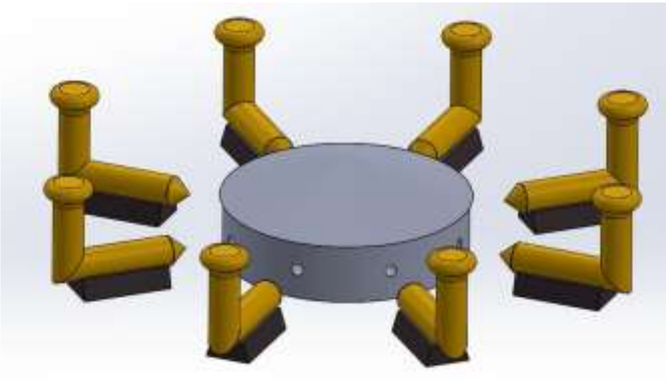
Nemanja Aranđelović, Dušan Nikezić, Dragan Brajović, Uzahir Ramadani

Apstract—U ovom radu se razmatra mogućnost merenja elektronskog impulsa iz elektronskog generatora za zagrevanje plazme elektrooptičkom metodom. Eksperimenti se vrše primenom Kerovog efekta na modelu elektronskog generatora. Kerov efekat pokazuje izuzetno dobre karakteristike za merenje impulsa nanosekundne brzine. Međutim, dobijeni rezultati pokazuju da te njegove karakteristike znatno kvari gama zračenje u dinamičkom stanju kao i apsorbovana doza gama zračenja. Kada se tome doda i jednostavnost merenja kapacitivnom sondom može se zaključiti da Kerov elektrooptički efekat nije preporučljiv za merenje u fuzionim eksperimentima.

Ključne reči— nuklearna fuzija, elektronsko zagrevanje plazme, Kerov elektrooptički efekat, uticaj polja gama zračenja.

I. UVOD

Rastuća potreba za energijom povećava napore za omogućavanje komercijalnog korišćenja nuklearne fuzije. Da bi do toga došlo potrebno je omogućiti zagrevanje plazme do temperatura koje su veće od praga za reakciju nuklearne fuzije. To se postiže injektovanjem energije u plazmu koja se nalazi u, takozvanoj, magnetnog boci. Prvobitno su rađeni eksperimenti sa laserskim injektovanjem energije u plazmu. Međutim, pokazalo se da plazma, pre postizanja temperature reakcije fuzije, počne da reflektuje laserske zrake čime se gubi veliki deo energije. Iz tog razloga se prešlo na čestično injektovanje energije u plazmu. To rešenje se pokazalo kao bolje, uz određene tehničke probleme koje treba rešiti [1-3].
Koncepcija čestičnog injektovanja energije u plazmu zasniva se na elektronskom generatoru, slika 1. Elektronski generator se sastoji od vertikalno postavljenog Marksovog generatora izolovanog uljem i horizontalnog talasovoda za oblikovanje elektronskog impulsa. Naime, elektronski generator treba da generiše elektronski impuls snage TW i širine 5 ns [4, 5]. Da bi se koncentrisalo dovoljno energije potrebno je da veći broj elektronskih generatora istovremeno injektuje elektronski impuls u plazmu. Iz zahteva za istovremenost injektovanja energije i nanosekunde širine impulsa javlja se problem džitera. Taj problem se rešava usavršavanjem okidanja Marksovog generatora primenom brzih troelektrodnih iskrišta [6-8]. Pored toga potrebno je meriti elektronski impuls na izlazu iz elektronskog generatora u cilju sinhronizacije okidanja svih generatora povratnom spregom. Merenje elektronskog snopa, zbog velike brzine, vrši se ili elektrooptičkim efektima ili brzim kapacitivnim sondama [9, 10].



Sl. 1: Osnovni oblik elektronskog generatora.

Pošto se elektronski generator sa pratećom opremom nalazi u intenzivnom polju gama zračenja može da dođe do njegovog uticaja na tačnost merenja elektronskog snopa. Iz tog razloga je cilj ovog rada da se proveri uticaj gama zračenja i doze gama zračenja na efikasnost merenja elektronskog snopa (tj. njegovog električnog polja) primenom elektrooptičkog Kerovog efekta.

II. KEROV ELEKTROOPTIČKI EFEKAT

Kerov elektrooptički efekat je pojava dvojnog prelamanja svetlosti u jakom električnom polju. Pošto je ovaj efekat otkriven eksperimentalno, opisuje se fenomenološki uvođenjem takozvane Kerove konstante. Kerova konstanta povezuje faznu razliku Δϕ između regularnog i neregularnog talasa dvojno prelomljenog svetlosnog talasa sa električnim poljem, tj.:

$$\Delta\varphi = 2\pi B l E^2 \quad (1)$$

gde je $B=k/\lambda$ Kerova konstanta, l je dužina puta svetlosti kroz dielektrik u kom se vrši dvojno prelamanje i E je jačina električnog polja [11-14].

Izraz 1 omogućava merenje vrednosti električnog polja (ili napona) merenjem fazne razlike između regularnog i neregularnog talasa. Osobine i prednosti merenja električnog polja primenom Kerovog efekta nad standardnim postupcima

Nemanja Aranđelović – Tehnološko-metalurški fakultet, Univerzitet u Beogradu, Karnegijeva 4, 11000 Beograd, Srbija (e-mail: snemi76@gmail.com).
Dušan Nikezić Institut za nuklearne nauke „Vinča“ - Institut od nacionalnog značaja za Republiku Srbiju, Univerzitet u Beogradu, Mike Petrovića Alasa bb., 1100 Beograd, Srbija (e-mail: dusan@vin.bg.ac.rs).
Dragan Brajović – Fakultet tehničkih nauka u Čačku, Univerzitet u Kragujevcu, Svetog Save 65, 32102 Čačak, Srbija (email: dragan.brajovic@vstss.com)
Uzahir Ramadani – Institut za nuklearne nauke „Vinča“ - Institut od nacionalnog značaja za Republiku Srbiju, Univerzitet u Beogradu, Mike Petrovića Alasa bb., 1100 Beograd, Srbija (e-mail: uzahir@vin.bg.ac.rs).

tehnike visokih napona su: 1- veoma širok opseg visokonaponskih veličina koje se mogu meriti sa malom mernom nesigurnošću (1 %); 2- malo kašnjenje (oko 10^{-12} sekundi) step funkcije električnog polja; 3- pri korišćenju pravougaonih impulsa može se primeniti ista logika koja se koristi za jednosmerno električno polje; 4- dinamički opseg napona koji se može izmeniti sa zahtevanom tačnošću znatno prevazilazi 1 : 1000 (što je odgovarajući opseg klasične opreme); 5- primena Kerovog efekta nema probleme sa indukovanim naponima i prenaponima u galvanskoj vezi; 6- mogućnost merenja visokonaponskih veličina ekstremno visokih učestanosti; 7- mogućnost direktnih merenja u izolacionim uljima i drugim tečnim izolatorima i 8- mogućnost veoma precizne sinhronizacije optičkog i električnog signala [15-17].

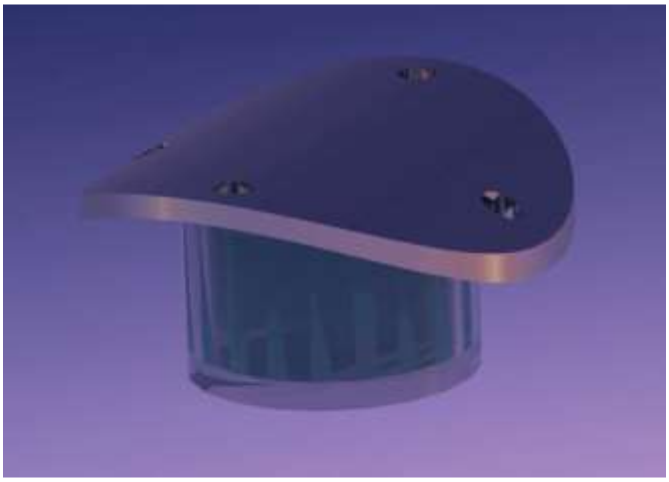
III. EKSPERIMENT I OBRADA EKSPERIMENTALNIH REZULTATA

Eksperimentalna provera dejstva gama zračenja i doze gama zračenja na mernu nesigurnost merenja pravougaonih naponskih impulsa vršena je na modelu horizontalnog dela elektronskog generatora, slika 2. U model sa slike 2 bili su ugrađeni polarizator, analizator i Kerova ćelija, slika 3. Kerova ćelija primenjivana u eksperimentu je bila zasnovana na tečnom dielektriku za dvojno prelamanje svetlosti. Primenjivane tečnosti i njihova Kerova konstanta su dati u tabeli 1.

Sl. 2. Model horizontalnog dela elektronskog generatora.

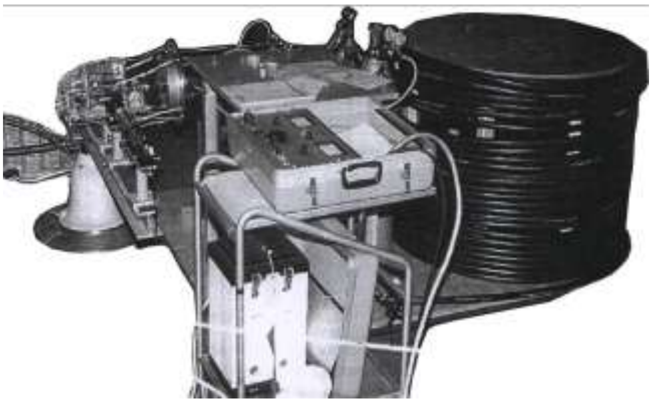
TABELA I
Kerove konstante primenjivanih tečnosti

TEČNOST	KEROVA KONSTANTA $[m/V^2]$
Nitrobenzen	$3.1 \cdot 10^{-12}$
Voda	$4.7 \cdot 10^{-14}$
Transformatorsko ulje	$1.8 \cdot 10^{-15}$



Sl. 3. Kerova ćelija koja se ugrađuje u horizontalni deo elektronskog generatora.

Kao izvor napona služio je visokonaponski kablovski generator, slika 4. Impuls iz kablovskog generatora je bio pravougaonog oblika sa veoma kratkim vremenom porasta i opadanja. Visokonaponski generator se sastojao od: 1- visokonaponskog transformatora; 2- ispravljača visokog naizmeničnog napona; 3- otpornika za ograničenje struje punjenja kabla; 4- koaksijalnog vodenog otpornika 50 Ω i 5- koaksijalnog visokonaponskog kabla dužine 60 m. Ovaj kablovski generator je generisao pravougaone impulse temene vrednosti 50 kV i trajanja 600 ns. Vremena uspona i opadanja čela i začelja su bila oko 3 ns.

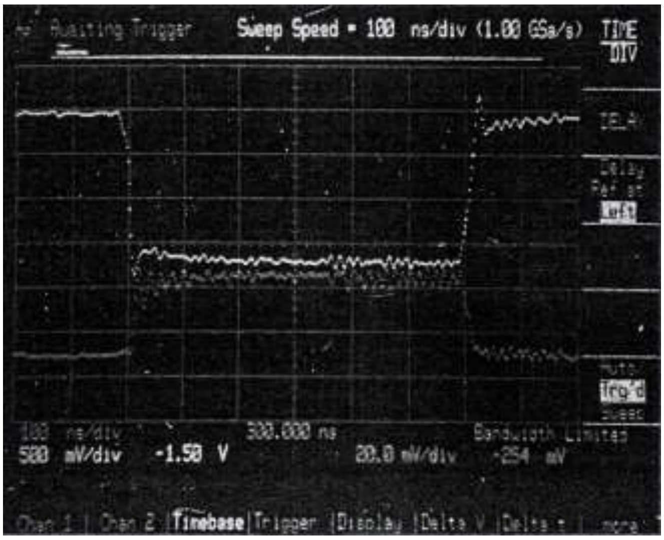


Slika 4: Visokonaponski kablovski generator.

Kao izvor svetlosti, tokom eksperimenta, korišćen je He-Ne laser. Ovaj tip lasera je jednostavan za rad i izvor je koherentne svetlosti. Osnovna talasna dužina ovog lasera je 632.82 nm.

Tokom merenja impulsa Kerovom ćelijom model horizontalnog dela elektronskog generatora je bio izložen dejstvu gama zračenja. Pri tome su vršena dva tipa merenja: 1- merenje pod dejstvom gama zračenja i 2- merenje nakon ozračenja Kerove ćelije pri čemu je data doza zračenja bila parametar merenja. Kao izvor zračenja korišćen je Co-60. Beta raspad radionuklida Co-60 prati fotonsko (gama) zračenje od 1.33 MeV i 1.17 MeV sa verovatnoćom emisije 1. Beta čestice koje se pri tome emituju nisu imale mogućnost da dopru do Kerove ćelije. Referentna merenja izvršena su u veličini kerma u vazduhu free-in-air pomoću referentnog mernog lanca, jonizacione komore i elektrometra, koji poseduje sledljivost do primarnog etalona. Prilikom dinamičnog merenja prototip horizontalnog dela elektronskog generatora postavljen je tako da je tokom merenja geometrija bila stalna. Doza koja je apsorbovana u Kerovoj ćeliji određivana je trajanjem ozračavanja na osnovu jačine doze određene jonizacionom komorom. Primenjene doze su iznosile: 20 Gy, 30 Gy, 40 Gy, 50 Gy, 60 Gy, 70 Gy, 80 Gy, 90 Gy, 100 Gy, 120 Gy, 140 Gy, 160 Gy, 180 Gy, 200 Gy, 300 Gy, 1000 Gy, 3000 Gy i 10000 Gy.

Merni postupak se sastojao u merenju impulsa iz kablovskog generatora Kerovom ćelijom. U postupku je vršena standardna konverzija fazne razlike regularnog i neregularnog talasa, slika 5 [18, 19]. Parametar merenja je bila vrsta dielektrika u Kerovoj ćeliji. Vršeno je po 50 snimanja napona iz kablovskog generatora dinamički i statički (za svaku dozu). Prilikom svakog merenja određivana je amplituda napona, vreme porasta, vreme opadanja i ripl na horizontalnom (temenom) delu napona. Takođe su statička merenja ponovljena (u istoj konfiguraciji) pet meseci nakon ozračenja. Merna nesigurnost tipa B postupka bila je 1.5 % [20, 21].



Sl. 5. Konvertovana fazna razlika regularnog i neregularnog talasa u naposnki impuls.

Statistička obrada eksperimentalnih rezultata se sastojala u sledećem: 1- primenom Šoveneovog kriterijuma čišćeni su statistički uzorci izmerenih slučajnih promenljivih od sumnjivih rezultata; 2- primenom χ^2 testa određena je statistička raspodela koja najbolje fituje dobijene statističke uzorke (testirane su Normalna raspodela i raspodela ekstremnih vrednosti); 3- momentnom metodom su određivana prva tri centralna momenta dobijenih raspodela i 4- određivana je merna nesigurnost tipa A za svaki statistički uzorak [22-24].

IV. Rezultati i diskusija

U tabeli 2 su date eksperimentalno dobijene vrednosti za merenje impulsa kablovskog generatora bez polja gama zračenja i sa poljem gama zračenja. Parametar eksperimentalnih rezultata su bile primenjene tečnosti u Kerovoj ćeliji.

TABELA II
EKSPERIMENTALNO DOBIJENE VREDNOSTI IMPULSA KABLOVSKOG GENERATORA BEZ I SA POLJEM GAMA ZRAČENJA

Tečnost	Amplituda		Vreme porasta		Vreme opadanja		Ripl		Merna nesigurnost tipa A	
Uslovi	bez γ	γ	bez γ	γ	bez γ	γ	bez γ	γ	bez γ	γ
Nitrobenzen	50 kV	46 kV	3 ns	3 ns	2 ns	2 ns	1 %	5 %	2 %	3.8 %
Voda	50 kV	41 kV	3 ns	3 ns	2 ns	2 ns	1.5 %	6 %	2.5 %	4.2 %
Trafo ulje	50 kV	43 kV	3 ns	3 ns	2 ns	2 ns	1 %	5 %	2.8 %	4.5 %

U tabeli 3 su date eksperimentalno dobijene vrednosti za merenje impulsa kablovskog generatora primenom Kerove ćelije ispunjene nitrobenzenom u zavisnosti od doze gama zračenja.

TABELA III
EKSPERIMENTALNO DOBIJENE VREDNOSTI IMPULSA KABLOVSKOG GENERATORA PRIMENOM KEROVE ĆELIJE ISPUNJENE NITROBENZENOM U ZAVISNOSTI OD PRIMLJENE DOZE GAMA ZRAČENJA

Doza [Gy]	Amplituda [kV]	Vreme porasta [ns]	Vreme opadanja [ns]	Ripl [%]	Merna nesigurnost tipa A [%]
20	50	3	2	1	2
30	50	3	2	1	2.2
40	49.3	3	2	1.2	2.2
50	48.8	3	2	1.4	2.25
70	47.8	3	2	1.4	2.32
80	46.4	3	2	1.52	2.37
90	45.8	3	2	1.6	2.46
100	45	3	2	1.65	2.62
120	44.1	3	2	1.76	2.84
140	43.2	3	2	1.88	2.93
160	42	3	2	2.2	3.1
180	40.8	3	2	2.4	3.4
200	39.4	3	2	2.7	3.7
300	35.2	3	2	3	4.5
1000	30	3	2	5.2	5.4
3000	27	3	2	6.8	5.6
10000	16	3	2	8	6.2

U tabeli 4 su date eksperimentalno dobijene vrednosti za merenje impulsa kablovskog generatora primenom Kerove ćelije ispunjene vodom u zavisnosti od doze gama zračenja.

TABELA IV
EKSPERIMENTALNO DOBIJENE VREDNOSTI IMPULSA KABLOVSKOG GENERATORA PRIMENOM KEROVE ĆELIJE ISPUNJENE VODOM U ZAVISNOSTI OD PRIMLJENE DOZE GAMA ZRAČENJA

Doza [Gy]	Amplituda [kV]	Vreme porasta [ns]	Vreme opadanja [ns]	Ripl [%]	Merna nesigurnost tipa A [%]
20	50	3	2	1	1.8
30	49.8	3	2	1.1	1.9
40	49.2	3	2	1.2	1.95
50	49.1	3	2	1.3	2.1
70	49.3	3	2	1.35	2.2
80	47.2	3	2	1.4	2.8
90	46.4	3	2	1.6	3
100	45.1	3	2	1.68	3.2
120	45	3	2	1.72	3.25
140	44.6	3	2	1.79	3.54
160	44.5	3	2	1.94	3.65
180	44.3	3	2	2.1	3.75
200	41.2	3	2	2.3	3.84
300	38.1	3	2	2.6	3.97
1000	33.8	3	2	2.9	4.2
3000	25.9	3	2	3.2	4.8
10000	16	3	2	5.4	5.9

U tabeli 5 su date eksperimentalno dobijene vrednosti za merenje impulsa kablovskog generatora primenom Kerove ćelije ispunjene trafo uljem u zavisnosti od doze gama zračenja.

TABELA V
EKSPERIMENTALNO DOBIJENE VREDNOSTI IMPULSA KABLOVSKOG GENERATORA PRIMENOM KEROVE ĆELIJE ISPUNJENE TRAFO ULJEM U ZAVISNOSTI OD PRIMLJENE DOZE GAMA ZRAČENJA

Doza [Gy]	Amplituda [kV]	Vreme porasta [ns]	Vreme opadanja [ns]	Ripl [%]	Merna nesigurnost tipa A [%]
20	49.8	3	2	1.8	2.1
30	49.8	3	2	2.1	2.15
40	48.2	3	2	2.3	2.25
50	48.1	3	2	2.35	2.38
70	47.8	3	2	2.39	2.41
80	47.6	3	2	2.41	2.43
90	47.1	3	2	2.43	2.46
100	47.0	3	2	2.47	2.49
120	46.8	3	2	2.49	2.52
140	46.1	3	2	2.52	2.54
160	46.0	3	2	2.55	2.59
180	45.7	3	2	2.59	3.12
200	45.1	3	2	3.2	3.21
300	44.9	3	2	3.6	3.28
1000	44.8	3	2	3.7	4.1
3000	44.0	3	2	4.1	4.9
10000	42.0	3	2	5.8	5.1

Na osnovu rezultata prikazanih u tabeli 2 se može zaključiti da Kerova ćelija u uslovima izloženosti dinamičkom gama zračenju menja neke od svojih karakteristika. Te promene su najviše izražene u slučaju da je tečnost u Kerovoj ćeliji trafo ulje, a najmanje izražene u slučaju da je tečnost u Kerovoj ćeliji nitrobenzen.

Dinamičko gama zračenje najviše utiče na ripl horizontalnog dela pravougaonog zračenja, a uopšte ne utiče na vreme porasta i opadanja pravougaonog impulsa. Pored toga dinamičko gama zračenje utiče na mernu nesigurnost tipa A merenja amplitude pravougaonog impulsa. Međutim, uticaj na ripl i amplitudu pravougaonog impulsa kablovskog generatora koje se pojavljuje kao posledica rada Kerove ćelije nisu značajne za sinhronizaciju rada elektronskih generatora. Osnovni uslov za sinhronizaciju elektronskih generatora je da vreme porasta i vreme opadanja impulsa budu konstantni, kao što jesu. Promene vrednosti amplitude pravougaonog impulsa, ripla i merne nesigurnosti tipa A takođe nisu od bitne važnosti za sinhronizaciju impulsa iz više elektronskih generatora pošto ne utiču na džiter. Što se tiče najvećeg uticaja gama zračenja na merenje Kerovom ćelijom kada je tečnost u ćeliji voda posledica je jednostavne strukture molekula vode i njene lake disocijacije.

Slična je situacija i sa primljenom dozom gama zračenja. Pošto primljena doza ne utiče na vreme porasta i opadanja impulsa može se zaključiti da primljena doza gama zračenja ne utiče na džiter. Ali prilikom velikih doza znatno opada amplituda pravougaonog impulsa, a to dovodi u pitanje ukupnu injektovanu energiju u plazmu. Najveći uticaj doza zračenja na amplitudu ima u slučaju da je Kerova ćelija ispunjena nitrobenzenom, a najmanji ako je Kerova ćelija ispunjena trafo uljem. To se može objasniti činjenicom što komponente nastale disocijacijom trafo ulja imaju približno istu vrednost Kerove konstante kao i trafo ulje. U slučaju nitrobenzena to nije slučaj. Manje opadanje amplitude sa povećanjem doze zračenja je proces rekombinacije atoma kiseonika i molekula vodonika tokom ozračivanja. Istim efektom se može objasniti i najveći stepen povratka prethodnih karakteristika Kerove ćelije sa vodom kao dielektrikom nakon vađenja iz polja gama zračenja. Naime, pokazalo se da je oporavak Kerove ćelije sa vodom kao dielektrikom nakon šest meseci neizloženosti gama zračenju stepena regeneracije 60 %. Regeneracija Kerove ćelije sa nitrobenzenom kao dielektrikom za isti period neizloženosti zračenju je samo 20 %. Ovako mali stepen regeneracije se objašnjava složenim molekulom nitrobenzena čije komponente disocijacije nemaju značajniju vrednost Kerove konstante. U slučaju Kerove ćelije punjene trafo uljem stepen regeneracije nakon šest meseci neizloženosti gama zračenju je oko 50 %. Međutim, velika vrednost ripla u ovom slučaju ukazuje da se ne radi o potpunoj regeneraciji trafo ulja već da u procentu oporavka učestvuju i komponente disocijacije koje imaju značajnu vrednost Kerove konstante.

Za razliku od ovih rezultata merenje uticaja neutronskog i gama zračenja na kapacitivnu sondu sa talasovodnim završnim otporom koja se takođe koristi za merenje impulsa elektronskog generatora pokazala je zanemarljiv efekat [25-27].

V. ZAKLJUČAK

Prethodno navedene osobine Kerovog elektrooptičkog efekta ga čine idealnim za merenje brzih naponskih impulsa. Međutim, rezultati prikazani u ovom radu pokazuju da Kerov efekat nije pogodan za merenje ako je Kerova ćelija ispunjena uobičajenim tečnostima, a treba da radi u polju gama zračenja. Naime, pokazuje se da gama zračenje ne utiče na regularni talas dvojnog prelamanja pošto njegovo prostiranje ne zavisi od indeksa prelamanja. Međutim, neregularna komponenta dvojnog prelamanja zavisi od indeksa prelamanja (tačnije njen indeks prelamanja zavisi od ugla upadnog zraka sa optičkom osom). Na taj način pravac kretanja neregularnog talasa postaje zavistan od indeksa prelamanja, a indeks prelamanja se menja usled pojave disocijacije molekula tečnosti usled zračenja. Na taj način dolazi do promenljive vrednosti fazne razlike između regularnog i neregularnog talasa. Ta razlika rezultira promenom amplitude izmerenog impulsa, njegovog ripla i merne nesigurnosti tipa A. Takvo ponašanje Kerove ćelije u polju gama zračenja čini je nepodesnom za merenje izlaznog elektronskog snopa elektronskog generatora koji se koristi za zagrevanje plazme. Ako se ovim osobinama doda i komplikovanost aparature za elektrooptička merenja može se zaključiti da je za ovaj tip merenja mnogo povoljnije koristiti brzu kapacitivnu sondu koja se pokazala rezistentna na dejstvo neutronskog i gama zračenja i koja daje zadovoljavajuće rezultate uz znatno jednostavniji merni sistem.

LITERATURA

[1] McCracken, G., Stott, P., Fusion, (2013) Fusion, 228 p, Elsevier Inc.

[2] Herman, R., (1990) Fusion: The Search for Endless Energy, Cambridge University Press, New York, NY. 527 pages. ISBN: 0-521-38373-0.

[3] Stankovic, K., Osmokrovic, P., Dolicanin, C., Vujisic, M., Vasic, A., Time enlargement law for gas pulse breakdown, (2009) Plasma Sources Science and Technology, 18 (2), art. no. 025028.

[4] Osmokrovic, P., Kusic, D., ArSic, N., Lazarevic, Z., Numerical and experimental design of three-electrode spark gap for synthetic test circuits, (1994) IEEE Transactions on Power Delivery, 9 (3), pp. 1444-1450.

[5] Osmokrović, P., Arsić, N., Vujisić, M., Stanković, K., Dolićanin, C., Reliability of three-electrode spark gaps, (2008) Plasma Devices and Operations, 16 (4), pp. 235-245.

[6] Avdic, S., Beganovic, A., Lagumdžija, A., Sadikovic, S., Pehlivanovic, B., Demirovic, D., Kadic, I., Comparative analysis of gamma dose rates measured by ion chamber in and around the historical sacral objects in Bosnia and Herzegovina, (2019) Nuclear Technology and Radiation Protection, 34 (2), pp. 157-164.

[7] Nedić, T.M., Janićijević, A.J., Stanković, K.D., Kartalović, N.M., Efficient replacement of the radioactive sources in the gas-filled surge arresters construction for the insulation co-ordination at the low voltage level, (2020) Nuclear Technology and Radiation Protection, 35 (2), pp. 130-137.

[8] S. B. Djekic, et al.: Passive and Active Shielding Against Electromagnetic Radiation, Nuclear Technology & Radiation Protection: Year 2020, Vol. 35, No. 4, pp. 331-338.

[9] D.C. Wunsch, *Kerr Cell Measuring System for High Voltage Pulses*, The Review of Scientific Instruments, Vol. 35, No. 7, July 1964.

[10] S. Newton, *Kerr Cell System for the Measurement of High Voltage Waveforms*, J.S. Instruments, Vol. 44, 1967.

[11] C.J.F. Böttscher, *Theory of Electric Polarization*, Dielectric in Static Fields, Vol. I, Elsevier, Amsterdam, 1973.

[12] Trifunović-Dragišić, V.Z., Stanković, M.D., Brajović, D.V., Kartalović, N.M., Estimation of the lifetime of solar cells in real conditions using accelerated aging under the influence of neutron and gamma radiation, (2019) Nuclear Technology and Radiation Protection, 34 (3), pp. 256-263.

[13] Fan, X., Chen, W., Zhang, Q., Chen, M., Wen, T., Zhang, X., Pei, Z., He, T., An electric field measurement method based on electro-optical modulation for corona discharge in air, (2019) Review of Scientific Instruments, 90 (8), art. no. 084704.

[14] Arbutina, D.S., Vasić-Milovanović, A.I., Nedić, T.M., Janićijević, A.J., Timotijević, L.B., Possibility of achieving an acceptable response rate of gas-filled surge arresters by substitution of alpha radiation sources by selection of electrode material and the electrode surface topography, (2020) Nuclear Technology and Radiation Protection, 35 (3), pp. 223-234.

[15] P. Osmokrović, I. Milovanović, M. Vujisić, K. Stanković, R. Radosavljević, „Experimental measurements of very fast transient voltages based on an electro-optic effect“, International Journal of Electrical Power & Energy Systems, Vol. 43, No. 1, pp. 408 – 417, 2012.

[16] Stanojević, N.Z., Djokić, J.V., Osmokrović, P.V., Research on water interconnections within the Šar mountains aquatorium by radioactive hydrogen isotope tritium, (2019) Nuclear Technology and Radiation Protection, 34 (4), pp. 392-398.

[17] H.M. Hertz, "Kerr effect tomography for nonintrusive spatially resolved measurements of asymmetric field distributions", Applied Optics, Vol. 25, No. 6, pp. 914 – 921, March 1986.

[18] P. Osmokrović, D. Filipović, M. Pešić and Z. Lazarević, „Transient electric field measurement in the liquid dielectrics using computerized laser tomography“, *IEEE Transactions on Instrumentation and Measurement*, Vol. 56, No. 6, pp. 2538 – 2546, December 2007.

[19] M. Stojkanović, G. Djukić, K. Stanković, M. Vujisić and P. Osmokrović, "Design Deployment and verification of the capacitive voltage divider for measuring fast transient occurrences in the nanosecond range“, Int. J. Elec Power, Vol. 43, pp. 1479-1486, 2012.

[20] Vulević, B., Osmokrović, P., Evaluation of uncertainty in the measurement of environmental electromagnetic fields, (2010) Radiation Protection Dosimetry, 141 (2), art. no. ncq158, pp. 173-177.

[21] P. Osmokrović, D. Petković and O. Marković, "Measuring probe for fast transients monitoring in gas - insulated substations“, ETEP/IEEE Trans. Instrum. Meas., Vol. 46, pp. 36-45, 1997.

[22] Lazarevic, Z., Radosavljevic, R., Osmokrovic, P., A novel approach for temperature estimation in squirrel-cage induction motor without sensors, (1999) IEEE Transactions on Instrumentation and Measurement, 48 (3), pp. 753-757.

[23] Kaur, R., Shikha, D., Singh, S.P., Mehta, V., Environmental radon, its exhalation rates and activity concentration of ^{226}Ra, ^{232}Th, and40k in Northern India, (2020) Nuclear Technology and Radiation Protection, 35 (3), pp. 268-282.

[24] Polužanski, V.S., Žigić, A.D., Kovačević, D.S., Nikolić, B.D., Computer non-iterative data acquisition of particle trajectory in a spark chamber, (2019) Nuclear Technology and Radiation Protection, 34 (1), pp. 65-71.

[25] Kartalović, N., Stanković, K., Nikezić, D., Kovačević, U., Dose effect of gamma radiation on reliable voltage pulse measurement in nuclear fusion experiments, (2020) Radiation Effects and Defects in Solids.

[26] Pejović, M.M., Pejović, M.M., Jakšić, A.B., Stanković, K.D., Marković, S.A., Successive gamma-ray irradiation and corresponding post-irradiation annealing of pMOS dosimeters, (2012) Nuclear Technology and Radiation Protection, 27 (4), pp. 341-345.

[27] Osmokrovic, P., Stojanovic, M., Loncar, B., Kartalovic, N., Krivokapic, I., Radioactive resistance of elements for over-voltage protection of low-voltage systems, (1998) Nuclear Instruments and Methods in Physics Research, Section B: Beam Interactions with Materials and Atoms, 140 (1-2), pp. 143-151.

ABSTRACT

In this paper, the possibility of measuring the electronic pulse from an electronic generator for plasma heating by electro-optical method is considered. The experiments are performed by applying the Kerr effect on an electronic generator model. Kerr effect shows very good characteristics for measuring nanosecond pulse rate. However, the results obtained show that these characteristics are significantly spoiled by gamma radiation in a dynamic state as well as by the absorbed dose of gamma radiation. When the simplicity of measuring with a capacitive probe is added to that, it can be concluded that the Kerr electro-optical effect is not recommended for measurement in fusion experiments.

Influence of gamma radiation on measurement fast pulse voltages by Kerr electro-optic effect

Nemanja Aranđelović, Dušan Nikezić, Dragan Brajović, Uzahir Ramadani

Radioactive Waste Management: Construction and Demolition Debris in Geopolymers

Ivana Jelić, Marija Šljivić-Ivanović, Tatjana Miljojčić, Milica Ćurčić and Slavko Dimović

Abstract — **Construction and demolition debris (C&DD) is one of the fastest-growing waste streams due to the global economic development and urbanization process. Therefore, developing more attractive and inexpensive methods and creating more valuable conventional and novel technologies that could more efficiently use these wastes and solve possible environmental problems, especially radioactive waste. The most widespread and economically viable solution for the reuse of C&DD today is civil engineering and the road industry. Also, there are several possible ways to use C&DD in geopolymers as recycled aggregates, activating components (precursors) depending on the composition, and as a hybrid system: with some aluminosilicate material that has better geopolymerization capacity or ordinary Portland cement. This use of C&D enables the synthesis of a wide range of matrices for the immobilization of radionuclides.**

Index Terms — **radionuclides; immobilization; geopolymers; environment; raw materials.**

I. Introduction

Construction and demolition waste (C&DW) is generated during the production of construction products or semi-final products, construction, demolition, and reconstruction. This waste accounts for the largest source of the solid waste stream in most countries worldwide [1-2]. According to European Environment Agency (EEA): “Construction and demolition waste (C&DW) comprises the largest waste stream in the EU, with relatively stable amounts produced over time and high recovery rates. Although this may suggest that the construction sector is highly circular, scrutiny of waste management practices reveals that C&DW recovery is largely based on backfilling operations and low-grade recovery, such as using recycled aggregates in road sub-bases” [3] as unbound aggregates or bound aggregates for concrete mixtures [4]. It is considered that construction is liable for climate changes (50%), increased energy consumption (40%), landfill waste generation (50%), air and water pollution, destruction of natural habitats, and negative impact on human health [5-7]. Thus, C&DW has an overall adverse environmental and economic impact, as well, on the construction sector, contributing to the emergence of additional costs. The main goal of C&DW management is to establish sustainable waste management, monitoring the quantities, types and composition of waste, waste generation, reduction of waste amount, separation and disposal of all types of construction waste, and its recycling and reuse.

Construction and demolition debris (C&DD), as one of the C&DW main fractions, occurs preferably after demolishing or reconstructing buildings (Fig. 1). It represents parts of walls, concrete, ceramic and roof tiles, carpentry, electrical parts, i.e. “discarded materials generally considered to be not water-soluble and non-hazardous in nature, including, but not limited to, steel, glass, brick, concrete, asphalt roofing material, pipe, gypsum wallboard, and lumber, from the construction or destruction of a structure as part of a construction or demolition project or from the renovation of a structure, and including rocks, soils, the tree remains, trees, and other vegetative matter that normally results from land clearing or land development operations for a construction project, including such debris from the construction of structures at a site remote from the construction or demolition project site” [8].

Fig. 1. Construction and demolition debris

This waste type is one of the fastest-growing waste streams due to the global economic development and urbanization process. The rapid progress of the market-oriented production

Ivana Jelić – Vinca Institute of Nuclear Sciences, University of Belgrade, P.O. Box 522, 11000 Belgrade, Serbia (e-mail: ivana.jelic@vin.bg.ac.rs).
Marija Šljivić-Ivanović – Vinca Institute of Nuclear Sciences, University of Belgrade, P.O. Box 522, 11000 Belgrade, Serbia (e-mail: marijasljivic@vin.bg.ac.rs).
Tatjana Miljojčić – Vinca Institute of Nuclear Sciences, University of Belgrade, P.O. Box 522, 11000 Belgrade, Serbia (e-mail: tatjana.miljojcic@vin.bg.ac.rs).
Milica Ćurčić– Vinca Institute of Nuclear Sciences, University of Belgrade, P.O. Box 522, 11000 Belgrade, Serbia (e-mail: milica.curcic@vin.bg.ac.rs).
Slavko Dimović – Vinca Institute of Nuclear Sciences, University of Belgrade, P.O. Box 522, 11000 Belgrade, Serbia (e-mail: sdimovic@vin.bg.ac.rs).

economy, industrialization, and population growth generate millions of tons of C&DD per year. This waste type counts the highest percentage of waste worldwide, approximately 75% [1].

Since C&DD represents a significant part of the waste stream, it is aimed for necessarily waste reduction or recycling. C&DD presents a considerable amount of building materials that could be reused or renew, avoiding non-renewable raw materials depletion in the construction sector. This significant recovery and recycling potential is lost via a lack of waste collection facilities or poor recycling practices. Therefore, it is necessary to develop more attractive and low-cost methods and create more valuable conventional and novel technologies that could more efficiently use these wastes and solve possible environmental problems. Resource efficiency and the circular economy concept play an essential role in environmental and economic policy, as well as so-called “6R” principles in sustainable supply chain design (reduce, reuse, recycle, recover, redesign and remanufacture) [9].

Clean C&DD, without plastic, metal, rubber, and wood, is recycled into quality construction material, most often in the form of (unbound) aggregates. C&DD utilization is also increasingly popular in scientific research since it represents a sustainable and environmental-friendly solution. Waste materials utilization reduces the exploitation of non-renewable natural resources and the exploitation of various forms of energy to synthesize or modify natural raw materials or artificial materials. Availability and cost-effectiveness are of great importance for C&DD multipurpose utilization.

Benefits of C&DD reuse & recycling:

- Cost-benefit – lower disposal fees, less need to purchase new materials;
- Natural resources conservation;
- Slowing down the rate at which landfills reach capacity;
- Reducing methane emissions created when landfilled materials break down.

In particular, nowadays, there is a market for aggregates derived from C&DD, such as road base materials, drainage structures, and other construction projects, but its utilization potential is still under-used. Many studies have dealt with C&DD application possibilities in order to develop acceptable utilization techniques. A promising alternative recycling option appears to be offered by alkali-activated materials, i.e. geopolymers, incorporating C&DD as inert aggregates or partially reactive materials. Since geopolymers were shown substantial flexibility in various industrial wastes and by-products utilization, the use of C&DD in these binders has been extensively investigated, with encouraging results.

According to data, the immobilization of radionuclides in the waste-based geopolymers was rarely investigated, unlike very comprehensive research on heavy metals [13]. However, the advantage of these waste-based materials represents the possibility of using any waste containing aluminosilicate, which could be dissolved in an alkaline solution to obtain a matrix for immobilization of radionuclides.

II. C&DD UTILIZATION WORLDWIDE

As mentioned, the most widespread and economically viable solution for the reuse of C&DD today is the use in civil engineering and the road industry. Globally, there is a strong tendency to use recycled building materials exclusively in the same industry rather than expanding to other sectors.

The global demand for natural aggregates production of concrete is projected to grow by an average of 8% per year by 2022 [10]. Developing countries drive a significant portion of this demand as a result of their rapid industrialization and urbanization growth. The utilization of recycled aggregates (Fig. 2) from C&DD (concrete, bricks, tiles, plastics, etc.) could significantly help conserve natural resources and reduce waste disposal, which gains environmental and economic benefits. Though the incorporation of recycled aggregates from C&DD remarkably decreases the environmental impact and carbon footprint of concrete, the utilization of these aggregates in construction activities in developing countries is yet limited. The main reason is the lack of certainty in the properties of concretes or other construction materials with C&DD recycled aggregates under exploitation conditions. Thus, the use of recycled aggregates, manufactured from recycled products, to replace virgin aggregates and contribute to sustainable construction needs to be encouraged [11-12].

Fig. 2. Recycled Concrete Aggregates [11]

III. C&DD IN GEOPOLYMERS

Ordinary Portland cement (OPC), as the primary ingredient commonly used to prepare concrete binders, requires intensive energy in its production and produces a considerable amount of carbon dioxide and greenhouse gases. It is estimated that the cement industry contributes approximately 5% of global environmental pollution through carbon dioxide emissions [14]. With an estimated annual growth of 4% in cement production [15], carbon dioxide emissions will increase and cause additional environmental burdens [16]. Contrarily, geopolymer is featured with low greenhouse-gas emissions, less energy consumption, and reuse of waste materials, which is critical to future sustainability [17].

The standard procedure for radionuclide immobilization today is transformation into stable insoluble forms by matrix materials (solidification), most often in cement-based matrices.

The term geopolymer and its description as cement-free green cementitious material was coined nearly three decades back by Davidovits for alumino-silicate polymers formed in the alkaline environment [18]. Geopolymerization technology has been shown advantages in reusing various types of waste to produce new materials for many purposes. These so-called inorganic polymers have been proposed to utilize solid aluminosilicate waste and the development of new cementitious materials that could be made without OPC [19]. It is a novel family of building materials, a new material for coatings and adhesives, new binders for fiber composites, waste encapsulation, and new cement for concrete [20].

Geopolymers have attracted attention due to the simplicity of synthesis with low or zero greenhouse gas emissions [10,12]. Hence, the utilization of waste-based geopolymers could show many advantages such as usage of low-cost materials in production, e.g. slags, fly ash, clays, saving natural resources, ambient temperature production, and high compressive and flexural strengths, in particular as compared to cement [18-19]. All these characteristics are placing geopolymers in a category of eco-friendly and sustainable materials.

The chemical composition of geopolymer is somewhat similar to zeolites but with an amorphous microstructure. Unlike OPC/pozzolanic cement, geopolymers do not form calcium-silicate-hydrates (C-S-H) for matrix formation and strength but utilize the poly-condensation of silica and alumina precursors to attain structural strength. Davidovits [21] elucidated a structural model of the geopolymer and assumed an essentially monolithic polymer similar to organic polymers. (Fig. 3).

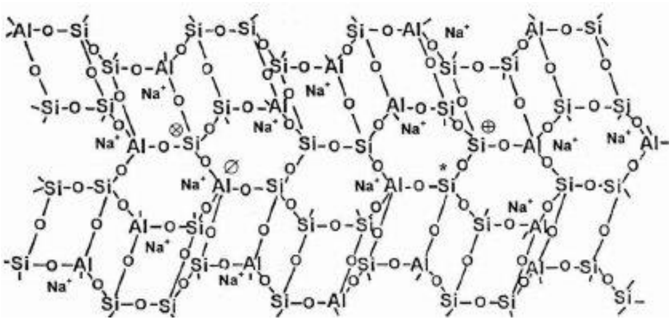


Fig. 3. Davidovits model of geopolymer structure [21]

Scientific publications provide a wealth of information relevant to productive C&DD usage in geopolymerization as processing technologies for value-added products.

Geopolymers based on C&DD differ in debris composition. Since its composition varies, there are many potential types of geopolymeric structures.

Fig. 4. Fly ash-based geopolymers [22]

Hence, there are several possible ways to use C&DD in geopolymers, as:

- recycled aggregates [4,23],
- activating component (precursors) [24-25] depending on the composition (e.g. bricks), and as
- hybrid system: with some aluminosilicate material with better geopolymerization capacity (e.g. metakaolin) or OPC [26].

The binary systems with OPC are synthesized due to the increase in geopolymer compressive strength in cases when it is assumed that geopolymerization will not be enough [26]. For example, this is the case with concrete debris, which is considered not to geopolymerize well due to its predominant participation of calcium carbonate in composition. Namely, the lack of aluminosilicate in almost pure concrete debris would lead to geopolymer non-creation. It was concluded that the addition of OPC in the geopolymerization processes contributed to the better mechanical behavior observed in the hybrid and binary systems. Although the use of cement is unfavorable from an energy point of view, this can affect the more significant usage of debris whose quantities are large, but geopolymerization is low.

However, it is more common for C&DD-based geopolymers to be synthesized with other aluminosilicates, especially waste-based material such as fly ash (Fig. 4) or other industrial waste: furnace slag, red mud, coal slag, etc. [4,27-29].

The stated utilization of C&DD allows the synthesis of a wide range of matrices for the immobilization of radionuclide.

IV. Conclusion

Recycling C&DD into sustainable and energy-efficient construction materials is a viable approach to relieve the stress of pollution and conserve virgin resources for the next generation. This study has reviewed the applications of C&DD as substitute materials in manufacturing sustainable geopolymer composites, specifically in the form of partial or even complete substitution of precursors or aggregates.

All results from cited studies suggest that waste-based geopolymers represent promising materials, but more thorough investigations are needed, especially for the utilization in radionuclide immobilization.

Acknowledgment

The research presented in this paper was done with the financial support of the Ministry of Education, Science and

Technological Development of the Republic of Serbia, within the funding of the scientific research work at the University of Belgrade, Vinca Institute of Nuclear Sciences (Contract No. 451-03-9/2021-14/ 200017).

REFERENCES

[1] V. A. Mymrin, K. P. Alekseev, R. E. Catai, R. L. S. Izzo, J. L. Rose, A. Nagalli, , C. A. Romano, "Construction material from construction and demolition debris and lime production wastes", *Constr. Build. Mater.*, vol. 79, pp. 207-213, 2015.

[2] O. Zimbili, W. Salim, M. A. Ndambuki, "Review on the usage of ceramic wastes in concrete production", *Int. J. Civ. Archit. Struct. Constr. Eng.*, vol. 8, pp. 91-95, 2014.

[3] EEA, Construction and demolition waste: challenges and opportunities in a circular economyhttps://www.eea.europa.eu/publications/construction-and-demolition-waste-challenges

[4] M. Panizza, M. Natali, E. Garbin, S. Tamburini, M. Secco, "Assessment of geopolymers with Construction and Demolition Waste (CDW) aggregates as a building material", *Constr. Build. Mater.*, vol. 181, pp. 119-133, 2018.

[5] M. Medineckienė, Z. Turskis, Z. Kazimieras, E. Zavadskas, "Sustainable Construction Taking into Account the Building Impact on the Environment", *J. Environ. Eng. Landsc. Manag.*, vol. 18, no. 2, pp. 118-127, 2010.

[6] L. Huang, G. Krigsvoll, F. Johansen, Y. Liu, X. Zhang, "Carbon emission of global construction sector", *Renew. Sust. Energ. Rev*., vol. 81, pp. 1906-1916, 2018.

[7] United States Environmental Protection Agency (EPA), National Overview: Facts and Figures on Materials, Wastes and Recycling, https://www.epa. gov/facts-and-figures-about-materials-waste-and-recycling/national-overview-facts-and-figures-materials

[8] C&D Debris Recycling Study: Final Report, https://floridadep.gov/sites/default/files/CDreport5-8.pdf

[9] S. Sakai, H. Yoshida, Y. Hirai, M. Asari, H. Takigami, S. Takahashi, K. Tomoda, M. V. Peeler, J. Wejchert, T. Schmidt-Unterseh, A. Ravazzi Douvan, R. Hathaway, L. D. Hylander, C. Fischer, J. G. Oh, L. Jinhui, N. C. Chi, "International comparative study of 3R and waste management policy developments", *J. Mater. Cycles Waste Manag*., vol. 13, pp. 86-102, 2011.

[10] M. C. Collivignarelli, G. Cillari, P. Ricciardi, M. C. Miino, V. Torretta, E. C. Rada, A. Abbà, "The Production of Sustainable Concrete with the Use of Alternative Aggregates: A Review", *Sustainability*, vol. 12, no. 7903, 2020.

[11] K. McNeil, T. H. K. Kang, "Recycled Concrete Aggregates: A Review", *Int J Concr Struct Mater*, vol. 7, pp. 61-69, 2013.

[12] A. Noorul, M. Faisal, K. Muhammad, S. Gul, "Synthesis and characterization of geopolymer from bagasse bottom ash, waste of sugar industries and naturally available China clay", *J. Clean. Prod*., vol. 129, pp. 491-495, 2016.

[13] T. Cheng, M. Lee, M. Ko, T. Ueng, S. Yang, S. The heavy metal adsorption characteristics on metakaolin-based geopolymer. Appl. Clay Sci., vol. 56, pp. 90-96, 2012.

[14] R. M. Andrew, "Global CO_2 emissions from cement production", *Earth Syst. Sci. Data*, vol. 10, pp.195–217, 2018.

[15] G. H. Kusuma, J. Budidarmawan, A. Susilowati, "Impact of Concrete Quality on Sustainability", *Procedia Eng*., vol. 125, pp. 754-759, 2015.

[16] S. Dadsetan, H. Siad, M. Lachemi, M. Sahmaran, "Construction and demolition waste in geopolymer concrete technology: a review" *Mag. Concr. Res*., vol. 71, no. 23, pp. 1232-1252, 2019.

[17] Z. Tang, W. Li, V. W. Y. Tam, C. Xue, "Advanced progress in recycling municipal and construction solid wastes for manufacturing sustainable construction materials", *Resour Conserv Recy*, vol. 6, no. 100036, 2020.

[18] J. Davidovits, "Geopolymers", *J. Thermal Anal*., vol. 37, pp. 1633-1656, 1991.

[19] J. L. Provis, "Geopolymers and other alkali activated materials: why, how, and what?", *Mater.Struct*., vol. 47, pp. 11-25, 2014.

[20] R. Cioffi, L. Maffucci and L. Santoro "Optimization of geopolymer synthesis by calcinations and polycondensation of a kaolinitic residue Resources", *Resour Conserv Recycl*, vol. 40, pp. 27-38, 2003.

[21] J. Davidovits, "Solid phase synthesis of a mineral block polymer by low temperature polycondensation of aluminosilicate polymers", IUPAC International Symposium on Macromolecules Stockholm; Sept. Topic III, New Polymers of high stability, 1976.

[22] S. Thokchom, P. Ghosh, S. Ghosh, "Performance of Fly ash Based Geopolymer Mortars in Sulphate Solution", J. Eng. Sci. Technol, vol. 3, no. 1, pp. 36-40, 2010.

[23] Z. Abdollahnejad, M. Mastali, M. Falah, T. Luukkonen, M. Mazari, M. Illikainen, "Construction and Demolition Waste as Recycled Aggregates in Alkali-Activated Concretes", *Materials*, vol.12, no. 4016, 2019.

[24] O. Şahin, H. İlcan, A. T. Ateşli, A. Kul, G. Yıldırım, M. Şahmaran, Construction and demolition waste-based geopolymers suited for use in 3-dimensional additive manufacturing", *Cem Concr Compos*, vol. 121, no. 104088, 2021.

[25] D. Kioupis, A. Skaropoulou, S. Tsivilis, G. Kakali, "Valorization of Brick and Glass CDWs for the Development of Geopolymers Containing More Than 80% of Wastes", *Minerals,* vol.10, no. 672, 2020.

[26] A. Vásquez, V. Cárdenas, R. A. Robayo, R. M. de Gutiérrez "Geopolymer based on concrete demolition waste", Adv Powder Technol, vol. 27, no. 4, pp. 1173-1179, 2016.

[27] N. M. Azad, S. M. Samindi, M. K. Samarakoon, "Utilization of Industrial By-Products/Waste to Manufacture Geopolymer Cement/Concrete", *Sustainability*, vol. 13, no. 2, pp. 873-879, 2021.

[28] S. R. Zedan, M. R. Maha, A. A. Doa, M. H. Aya, "Effect of demolition/construction wastes on the properties of alkali activated slag cement", *HBRC*, pp. 331-336, 2017.

[29] B. A. Ionescu, A Lzrescu, "A Review Regarding the Use of Natural and Industrial by-products in the Production of Geopolymer Binders", *Mater. Sci. Eng*., vol. 877, no. 012033, 2020.

A Strategic Means of Hybrid Warfare

Milica Ćurčić, Slavko Dimović, Ivan Lazović

***Abstract*—The modern security environment is undergoing a profound transformation. This transformation has been shaped by the emergence of new patterns of conflict and cooperation among state and non-state actors as well as the spread of globalization and new technologies. Also, the development of a new breed occurred, characterized by a combination of warfare methods and usage of different means of warfare. In the constellation of new wars, a hybrid wars stand out as a war that combines different strategies of warfare to achieve synergistic effects. The aim of the article is to analyze and describe characteristic of both non-state and state hybrid warfare, as well as the key elements that constitute strategic means of hybrid warfare. The usage of information weapons, cyber sphere and psychological means, in combination with conventional weapons of war, become main features of modern conflict. Modern technologies are the main factor that influenced and transformed warfare and their usage permeates every activity in hybrid war.**

***Index Terms* — Hybrid warfare, war, means of warfare, security, strategy, new technologies**

I. Introduction

Clausewitz's observation that war is "a merely continuation of policy by other means" [1] has become an incontestable maxim among security experts. During history, war occurred among centralized, hierarchically ordered, territorialized states in which big armies confronted each other on the battlefield, using similar strategies, tactics, and weapons. The war is still present in international relations, but with different forms and characteristics.

The nature of warfare is changing and as Williams noticed, there are three issues important to discuss: how useful is the concept of 'total war' for thinking about developments in warfare? What is the relationship between war and globalization; specifically, has globalization given rise to a 'new' type of warfare? What changes can be identified in the way advanced industrialized democracies in the West are waging war today compared to earlier historical periods? [2]

The post-cold war period is characterized by the emergence of non-state actors, new threats that are combined with globalization factors and the usage of sophisticated technology. In order to find a new analytical framework for understanding the nature of modern wars, new paradigms had been constructed. Those concepts are seeking to understand an ambiguous border between war and peace because from the perspective of modern wars, wartime, peacetime and crisis do not exist as separate phases [3]. In a more practical way, the main dilemma arises from questions of how it is possible for a weaker actor to win a war against a significantly stronger and more powerful enemy and how a stronger actor should oppose to a weaker one in an asymmetric conflict.

One of the basic characteristics of modern wars is the use of new means of warfare that avoid predictability and linear military operations. Technological progress has conditioned changes in all spheres of life, including conflict, wars and military operations. The new technologies allow the possibility to achieve strategic goals by unconventional and cognitive effects (technologies of social influence and manipulation, cyber sphere, information weapon, possibilities of significant damage of control system of a state) [4]. Those technologies appear to be increasingly adaptive and sophisticated, able to outpace state-based militaries in the dialectic and competitive learning cycle inherent to wars [5]. Technological advancements have furthered weapons and platform development, but also introduced new capabilities and vulnerabilities in the security arena, that additionally increase the complexities of contemporary conflicts.

II. Hybrid War as a New Form of Warfare

The security architecture of the modern world focuses on threats such as terrorism and radicalization, nonproliferation of WMD, securitization of migration, cyber and ecological threats etc. Most of these threats are dominantly posed by different non-state actors. Also security agendas introduce new types of wars that cannot be defined as conventional, traditional or classical wars. In order to clarify the different types of war in a contemporary security environment, as well as their basic characteristics, a number of scholars have attempted to define new types of war. In the literature we can find terms Unconventional war [6-8], Irregular war [5, 9], Fourth Generation of War [10-11], Unrestricted War [12] Compound War [13] and Asymmetric War [14-16]. All this approaches of modern war in different ways are pointing at the blurring of subjects, objects and dynamics in contemporary conflicts.

The following variables are most commonly used to determine the characteristics of modern wars: the main protagonists and units of analysis of war - states or non-state actors; the primary motives of actors (ideology, territorial secession, religion etc.); the spatial ranges: interstate, regional, or global; the technological means of violence – weapons and strategies of war; the social, material, and human impact of conflict, including patterns of human victimization and forced

Milica Ćurčić – Department of Physical Chemistry, "Vinca" Institute of Nuclear Sciences - National Institute of Republic of Serbia, University of Belgrade, Belgrade, Serbia (e-mail: milica.curcic@vin.bg.ac.rs).

Slavko Dimović – Department of Radiation and Environmental Protection "Vinca" Institute of Nuclear Sciences - National Institute of Republic of Serbia, University of Belgrade, Belgrade, Serbia (e-mail: sdimovic@vin.bg.ac.rs).

Ivan Lazović – "Vinca" Institute of Nuclear Sciences - National Institute of Republic of Serbia, University of Belgrade, Belgrade, Serbia (e-mail: ivan.lazovic@vin.bg.ac.rs).

human displacement; the influence of the political, economic and social structure on conflict [17].

The „hybrid war" emerged as the newest kind of war during the first decade of the 21^{st} century by key scholars who focused on "the blending or blurring character of conflict" [18]. This term was first used in 2006 to describe the strategy that Hezbollah used in the Lebanon war. Mixing an organized political movement with decentralized armed cells employing adaptive tactics in ungoverned zones, Hezbollah affirms an emerging trend. Highly disciplined, well-trained, distributed cells can contest modern conventional forces with an admixture of guerrilla tactics and technology in densely packed urban centers [19]. In this conflict Hezbollah conducted several technological surprises on Israeli forces. Hezbollah's fighters bypassed the complex surveillance system used by Israel to monitor its border with Lebanon led to the kidnapping of two Israeli soldiers and the killing of eight. The firing of a Noor anti-ship cruise missile (an Iranian version of the Chinese C-802) resulted in the loss of four Israeli sailors and the crippling of an Israeli missile ship. Two Merkava IV tanks were destroyed and their crews killed or wounded, probably by a combination of Raad anti-tank missiles (the Iranian version of the Russian Sagger AT-3) and advanced improvised explosive devices (IEDs) [20].

This case demonstrated the ability of a non-state actor to deconstruct the vulnerability of not only a powerful state, but Western-style militaries [21]. The increased number of actors, who innovatively combine different models of war, capacities and weapons in order to achieve strategic goals, has created fertile ground for the introduction of concept which explains characteristics of modern warfare. In that manner, hybrid war is becoming the dominant discourse in discussions of modern warfare as well as accepted and promoted by politicians, military experts and theorists as the basic concept of modern military strategies [22].

The term 'hybrid warfare' was introduced in theory by Frank Hoffman, a former US Marine officer, to influence on ingrained and outdated beliefs in the US military about the utility of military force in the post-Cold War environment. In Hoffman's view, hybrid warfare was a suitable analytical construct to explain the success of a relatively weak opponent, such as the Taliban, Al Qaeda, or Hezbollah, against the vastly technologically and numerically superior militaries both in Afghanistan and Iraq and in the 2006 Lebanon war against Iraq and Israeli forces [23]. Hoffman define hybrid war as a type of war that "incorporate a range of different modes of warfare, including conventional capabilities, irregular tactics and formations, terrorist acts including indiscriminate violence and coercion and criminal disorder. Hybrid war can be conducted by both states and variety of non-state actors" [18].

What makes a war "hybrid" and divers from other modern war is coordinated fusion of different modes of warfare, both military (use of force) and non-military (violence, irregular tactics, criminal disorder, terrorist acts), to achieve synergistic effects in the physical and psychological dimensions of conflict within the main battle space [18].

The evolution of hybrid warfare has two phases so far. The first phase is called non-state hybrid war, as it involves the action of non-state actors that combines conventional forces, whose actions are regulated by the rules and norms of law and traditional military custom, with unconventional forces that conduct operations of guerrilla warfare, terrorist activities and criminal activities. Characteristics of this phase of hybrid war are: non-state exhibit increased levels of military sophistication as they successfully develop modern weapons systems (such as anti-ship missiles, UAVs), new technologies (cyber, secure communication, sophisticated command and control), and tactics (combined arms) that are traditionally considered to be outside range of such actors. Non-state actors expanded the battlefield beyond the purely military realm and show the growing importance of non-military tools by including elements of information warfare (e.g. controlling the battle of the narrative and online propaganda, recruitment and ideological mobilization [24].

The second phase of the evolution of the term hybrid war called state hybrid war begins with the Ukrainian crisis and the Russian annexation of Crimea in 2014. Russian operations demonstrate that hybrid warfare can be conducted with great success by state actors. The main characteristics of Russian operation in Crimea as a prototype of the second phase of hybrid warfare are:

- non-declaration of the state of war;
- non-contact clashes between highly maneuverable interspecific fighting groups;
- annihilation of the enemy's military and economic power by short-time precise strikes in the strategic military and civilian infrastructure;
- massive use of high-precision weapons and special operations, robotics, and weapons that use new physical principles (direct-energy weapons – lasers, shortwave radiation, etc);
- use of armed civilians;
- simultaneous strike on the enemy's units and facilities in all of the territory;
- simultaneous battle on land, air, sea, and in the informational space.
- use of asymmetric and indirect methods;
- management of troops in a unified informational sphere [25]. Usage of non-military means, especially the use of information surprised Ukraine and represented significant factors for the realization of Russian plans in Crimea and announced future trends in warfare.

This transformation of used means surprised even Hoffman, whose definition of hybrid warfare is limited to a combination of tactics related to violence and irregular way of warfare between state and non-state actors. His definition did not recognize non-violent and non-military instruments like diplomatic, economic and financial activities, subversive political acts such as the creation or secret use of trade unions and non-governmental organizations as a front of actions, or information and propaganda operations through the use of fake websites and newspaper articles [22].

As Figure 1 shows hybrid warfare differs from other types of war in their initiation and prosecution, involve various sphere of social action, employee different strategies and means. Hybrid warfare is directed towards the whole society with the aim of destabilization and polarization. In this type of war, not only the military weaknesses are essential but also those that only society can generate: ethnic tensions, weak and corrupt institutions, economic or energy dependence. Based on these weaknesses, hybrid war applies on the full spectrum of activities ranging from media propaganda to terrorism through irregular and unassumed warfare [26].

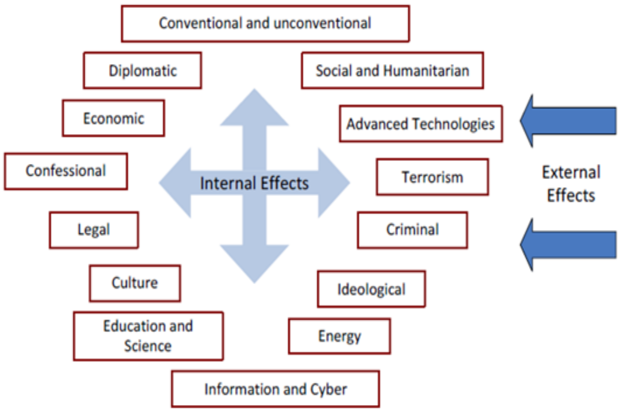


Fig 1. Hybrid warfare spheres [4]

III. The Impact of New Technologies on Hybrid Warfare

At the strategic level, the hybrid theory of warfare can be seen as the employment of information operations and diplomacy in conjunction with cyber and electronic operations to weaken an opponent or to sow the seeds of chaos in relation to an adversary [27]. In addition to traditional wars, hybrid wars are not declared and, therefore, cannot be completed in the classical sense of military conflicts. This is a permanent war of variable intensity across multiple sectors, with cascading impacts and synergistic destructive manifestations, in which the entire population of the country is involved. An essential feature of this concept is the diminished role of military content, more precise usage of armed struggle. Unlike the classic war conflicts, in which concepts based on the mass use of armed force were dominated, minimized often disguised military hard power is the most significant novelty in the history of warfare introduced by hybrid warfare [28].

Through the use of innovative technologies, it became possible to shift conflict from predominantly overt and forceful (kinetic) means to less obvious strategies focused on the structural vulnerabilities of adversaries, including achieving cognitive advantage over them [4]. Widespread usage of new technologies should provide reduction of hard military power to minimum creating a distorted image of the real attacker. In that way, modern technologies were the main factor that influenced and transformed warfare.

Hybrid war in Ukraine shows that the main battlefield is human mind so the most important elements of modern war become information and psychological means. Wide-ranging, multidimensional and by employing multifactorial information, hybrid warfare in Ukraine included applying of highly technological samples of weapons and military hardware. Y. Danyik et all in the paper [4] identifies some of the most important areas of information technology involved in hybrid warfare:

1. electronic warfare systems and complexes;
2. modern information and communication systems;
3. innovative weapon control systems;
4. integrated reconnaissance-strike complexes;
5. innovative software;
6. complexes for conducting information-psychological activities and actions in cyber space
7. environmental control and space systems;
8. robotic systems (especially unmanned aircraft complexes).

All modern country are highly depended on various information infrastructure and information-based resources including complex management systems and infrastructures involving the control of electric power, money flow, air traffic, oil and gas, and other information-dependent item [29]. The development and use of new, and especially information technologies is a determinant of the state development, but also the most important means in the application of measures and countermeasures in hybrid warfare.

Ukrainian hybrid war demonstrates the complexity of strategy that includes military and nonmilitary means relying on new technologies at every stage of operation. Moscow employed methods that blended conventional and irregular combat, economic coercion, sponsorship of political protests, and the now notorious disinformation campaign [30]. Also, different technologies were use simultaneously as a part of strategically design campaign with main goal of undermining public confidence in the government. Bērziņš, identified eight phases of Russian hybrid strategy:

1. non-military means (encompassing information, moral, psychological, ideological, diplomatic and economic measures);
2. special operations carried out by media, diplomatic channels, top government and military agencies to mislead political and military leaders (can include leaking false data, orders, directives, and instructions);
3. intimidating, deceiving and bribing government and military officers with the objective of making them abandon their service duties;
4. use of destabilizing propaganda to increase discontent among the population (can be further enhanced by the arrival of ‘volunteers’, escalating subversion);
5. establishment of no-fly zones over the targeted country, imposition of blockades, extensive use of private military contractors and armed opposition;
6. commencement of military action, immediately preceded by large-scale reconnaissance and subversive missions of all types (including special operations forces; space, radio, radio engineering, electronic, diplomatic and secret service intelligence; and industrial espionage);
7. targeted information, electronic warfare and aerospace operations along with continuous air-force harassment, combined with the use of high-precision weapons launched

from various platforms (long-range artillery and weapons based on new physical principles, including microwaves, radiation, radiological and ecological disasters and non-lethal biological weapons);

8. crushing the remaining points of resistance and destroying surviving enemy units by using special operation units. [25]

Those are phases of war that Russians refer as “new generation of warfare” directed against Western influence in the world. While the Chinese concept of ‘unrestricted warfare’ was aimed at identifying ways to counter the West’s over-whelming hard and soft power through asymmetric means, the Russians concept of warfare is the answer on tolls that Western use: liberalism, international institutions, non-governmental organizations, and strategic communication [30]. Hybrid warfare, or new generation of war demonstrated tremendous success by usage of a sophisticated blend of psychological warfare, cyber - attacks, strategic communication, disinformation campaign and covert troops. The further risks also arise from the circumstances that nuclear states do not directly confront each other by traditional means. The doctrinal turnover that includes strategic means of hybrid warfare, as well as military modernization of states, creates a new kind of security dilemma. In that sense, nuclear security based on the concept of nuclear deterrence should be reconsidered in the context of hybrid warfare. Hybrid warfare ignores a key concept that builds nuclear deterrence as a viable strategy including concepts of stability, preparedness, clarity and rationality.

IV. Conclusion

Warfare is sui generis a socio-historical phenomenon with a pronounced technological component. Definition of war modified with the change of social circumstances and due to technological progress. In the last two decades, this definition expanded to incorporate non-state actors, cyber warfare and usage of non-military means. The blending of all used means of waging the war is what distinguishes hybrid war from other historical forms. The mind becomes the main battlefield of the 21st century which puts focus on information and psychological means of warfare. In the further transformation of hybrid war, the tendency will be on in developing strategies and means how to first defeat adversary mentally by the usage of non-military means. The main goal of modern warfare is the reduction of hard military power and defeating the enemy in the short term without human losses, which hybrid warfare perfectly demonstrates.

Acknowledgment

The research was funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia, as the part of theme "Fundamental and applied aspects of electrocatalysis in energy systems and green technologies", within the program Energy and energy efficiency, which are realized in the “Vinča” Institute of Nuclear Sciences - National Institute of Republic of Serbia, University of Belgrade, Belgrade, Serbia.

References

[1] C. V. Clausewitz, “On War”, Oxford University Press, Oxford, 2007.
[2] P. D. Williams, “On War”, *In Security Studies – An Introduction.* Routledge, London and New York, pp. 151-170. 2008.
[3] G. Razma, “Modern Warfare Paradigm: Reconsideration of Combat Power Concept. Journal of Security and Sustainability Issues. Vol. 8, No.3, pp. 435-452. 2019.
[4] Y. Danyik, T. Maliarchuk, C. Briggs. “Hybrid War: High-tech, Information and Cyber Conflict”. *The Quartely Journal*, Vol. 10, No. 2, pp.5-24. 2017.
[5] F. G. Hoffman. “Complex Irregular Warfare: The Next Revolution in Military Affairs”. *Orbis*, Vol. 50, No. 3, pp. 395-411, 2006.
[6] YA. C. Janos, “Unconventional Warfare: Framework and Analysis”, *World Politics*, Vol. 15, No. 4, pp. 636-646, 1963.
[7] K. Knorr, “Unconventional Warfare : Strategy and Tactics in Internal Political Strife”, *The Annals of the American Academy of Political and Social Science*, Vol 341, Issue 1, pp. 53-64, 1962.
[8] D. Kilcullen, “The Evolution of Unconventional Warfare”. *Scandinavian Journal of Military Studies*, Vol. 2, No. 1, pp. 61–71. 2019.
[9] C. S. Gray, “Irregular Warfare: One Nature, Many Characters”, *Strategic Studies Quarterly*, Vol. 1, No. 2, pp. 35-57, 2007.
[10] W. S. Lind, K. Nightengale, J. F Schmitt, J. W. Sutton, Gary I. Wilso, “The Changing Face of War, Into the Four Generation“, *Marine Corps Gazette*, 1989.
[11] W. S. Lind, “Understanding Fourth Generation of War“, *Military Review*, 2004.
[12] Q. Liang, W. Xiangsui, *Unrestricted Warfare.* Beijing, PLA Literature and Arts 1999.
[13] T. Huber, “Compound Warfare: A Conceptual Framework”, *Compound Warfare: That Fatal Knot*, U.S. Army Command and General Staff College Press, Kansas, 2002, 1-11.
[14] D. Buffaloe, *Defining Asymmetric Warfare*, The Institute of Land Warfare, Virginia, 2006.
[15] A. Lele, “Asymmetric Warfare: A State vs Non-State Conflict”, *OASIS*, No. 20, 2014, 97-111.
[16] M. Ćurčić, Asymmetric Threats in Security Studies in *Asymmetry and Strategy - Thematic Collection of Articles*. Strategic Research Institute and National Defence School, Belgrade, pp. 17-31, 2018.
[17] E. Newman, “The ‘New Wars’ Debate: A Historical Perspective Is Needed. *Security Dialogue*. Vol. 35, No. 2, pp. 173–189. 2004.
[18] F. G. Hoffman, Conflict in the 21st Century – The Rise of Hybrid Wars, Potomac Institute for Policy Studies, Arlington, Virginia, 2007
[19] F. G. Hoffman, “Lessons from Lebanon: Hezbollah and Hybrid Wars”, Foreign Policy Research Institute, E-notes, 2006. Internet: https://www.fpri.org article/2006/08/lessons-from-lebanon-hezbollah-and-hybrid-wars/ Access: 29.05.2021
[20] G. Ben-Ari, “*Technological Surprise and Technological Failure in the Current Lebanon Crisis*”. Center for Strategic and International Studies. Wachington D.C. 2006.
[21] M. Vračar, M. Ćurčić, “The Evolution of European Perception of the term „Hybrid Warfare“, *Vojno delo*, Vol. 1, pp. 5-21. 2018.
[22] M. Ćurčić, “Hibridni ratovi u savremenom bezbednosnom okruženju“ *U Hibridno ratovanje - dilema koncepta savremenh sukoba*, Beograd, Srbija, 2018, 47-62.
[23] B. Renz, “Russia and ‘hybrid warfare”. *Contemporary Politics*. pp. 1-18. 2016.
[24] E. Reichborn-Kjennerud, P. Cullen, *What is Hybrid Warfare*, Policy Brief No.1. Norwegian Institute of International Affairs, 2016.
[25] J. Berzins, J. *Russia's New Generation Warfare in Ukraine: Implications for Latvian Defense Policy*, Policy Paper No. 2, Riga, National Defense Academy of Latvia, 2014.
[26] M. M. Neag, “A New Typology of War – The Hybrid War”, *Revista Academiei Forţelor Terestre*, No. 1, pp. 15-20, 2016.
[27] A. C. Fox. *Hybrid Warfare: The 21st Century Russian Way of Warfare*. School of Advanced Military Studies United States Army Command and General Staff College Fort Leavenworth. Kansas. 2017.

[28] M. Vračar, “Razmatranje adekvatnog teorijsko-epistemološkog pristupa u istraživanju fenomena hibridnog ratovanja“. *Vojno delo*, No. 7, str. 292-307. 2017.

[29] R. C. Molander, A. S. Riddile, P. A. Wilson, *Strategic Information Warfare – A New Face of War*. National Defense Research Institute. 1996.

[30] S. Hadjitodorov, M. Sokolov, “Blending New-generation Warfare and Soft Power: Hybrid Dimensions of Russia-Bulgaria Relations“. *Connections: The Quarterly Journal*. No. 1, pp. 5-20, 2018.

Standard and validated method for determination of tritium on Liquid scintillation spectrometer

Marija Janković, Nataša Sarap, Gordana Pantelić, Jelena Krneta Nikolić, Milica Rajačić, Dragana Todorović, Ivana Vukanac

***Abstract*—Tritium concentrations in water samples were analyzed. The aim is to compare methods: standard method (direct method): according to ASTM D 4107-08 and validated method: which applies electrolytic enrichment. Low level tritium concentration in natural waters required measurement after enrichment. One set of samples, which contain 16 samples including spike water were enrichment and compared with 4 samples measured by direct method. For this analysis liquid scintillation spectrometer was using. In general, analysis with enrichment is more applicable for samples with low activity. Also, validated method with enrichment reduced minimum detectable concentration.**

***Index Terms*—tritium, liquid scintillation spectrometer, water samples, enrichment.**

I. Introduction

Tritium (^3H), radioactive isotope of hydrogen, has half life of 12.3 years. ^3H has low beta energy with maximum of 18 keV and a mean energy of 5.7 keV. Its origin is naturally, from the upper atmosphere (stratosphere) where arises by interaction between fast neutron and nitrogen atoms, and artificially (produced from the nuclear reactor, atmospheric thermonuclear tests, nuclear weapons testing, nuclear reactors, fuel reprocessing plants, heavy water production facilities and commercial production for medical diagnostics, radiopharmaceuticals, luminous paints, sign illumination, self-luminous aircraft, airport runway lights, luminous dials, gauges and wrist watches) [1]. It can enter in hydrological cycle from precipitation and can enter in surface water and groundwater. From that reason, monitoring of tritium in natural water is necessary.

In water samples, tritium can be measured directly, after distillation, or after electrolytic enrichment of samples. Based on the fact that nowadays, concentration of tritium in natural water is low, it is necessary to concentrate the samples through enrichment. In the case of the assumption that the concentration is high, enrichment is not required.

This paper presents the comparison between different preparations of water samples which means using of standard method – direct method and validated method – method with electrolytic enrichment.

Marija Janković, University of Belgrade, Radiation and Environmental Protection Department, Vinča Institute of Nuclear Sciences, National Institute of the Republic of Serbia, University of Belgrade, Mike Petrovića Alasa 12-14, 11001 Belgrade, Serbia (e-mail: marijam@vinca.rs).

Nataša Sarap, University of Belgrade, Radiation and Environmental Protection Department, Vinča Institute of Nuclear Sciences, National Institute of the Republic of Serbia, University of Belgrade, Mike Petrovića Alasa 12-14, 11001 Belgrade, Serbia (e-mail: natasas@vinca.rs).

Gordana Pantelić, University of Belgrade, Radiation and Environmental Protection Department, Vinča Institute of Nuclear Sciences, National Institute of the Republic of Serbia, University of Belgrade, Mike Petrovića Alasa 12-14, 11001 Belgrade, Serbia (e-mail: pantelic@vinca.rs).

Jelena Krneta Nikolić, University of Belgrade, Radiation and Environmental Protection Department, Vinča Institute of Nuclear Sciences, National Institute of the Republic of Serbia, University of Belgrade, Mike Petrovića Alasa 12-14, 11001 Belgrade, Serbia (e-mail: jnikolic@vinca.rs).

Milica Rajačić, University of Belgrade, Radiation and Environmental Protection Department, National Institute of the Republic of Serbia, University of Belgrade, Vinča Institute of Nuclear Sciences, Mike Petrovića Alasa 12-14, 11001 Belgrade, Serbia (e-mail: milica100@vinca.rs).

Dragana Todorović, University of Belgrade, Radiation and Environmental Protection Department, Vinča Institute of Nuclear Sciences, National Institute of the Republic of Serbia, University of Belgrade, Mike Petrovića Alasa 12-14, 11001 Belgrade, Serbia (e-mail: beba@vinca.rs).

Ivana Vukanac, University of Belgrade, Radiation and Environmental Protection Department, Vinča Institute of Nuclear Sciences, National Institute of the Republic of Serbia, University of Belgrade, Mike Petrovića Alasa 12-14, 11001 Belgrade, Serbia (e-mail: vukanac@vinca.rs).

II. The Method

Total number of 15 water samples were collected in order to analyze tritium concentrations: 1 drinking water, November 2020, 9 precipitations (3 from referent meteorological station Zeleno Brdo in Belgrade (ZB), September, October and November 2020; 3 from station in Vinča Institute of Nuclear Sciences Site Center (CS), September, October and November 2020; 3 from station in Vinča Institute of Nuclear Sciences Meteo stub (MS), September, October and November 2020), 5 surface water (Sava, September and October 2020; Danube, October 2020; Mlaka Creek in Vinča Institute of Nuclear Sciences, October and November 2020).

Approximately 300 ml of each composite monthly sample are distilled to remove dissolved salts and organic impurities. The various ions that are naturally found in water (Cl^-, SO^{-4}, CO_3, Mg^{++}, Na^+, K^+, etc.) could interfere with the electrolysis process [2].

In the case where the direct method is applied the ASTM D 4107-08 standard method [3] is used. After preliminary distillation an aliquot of 8 ml of distilled samples is mix with 12 ml scintillation cocktail ULTIMA GOLD LLT in polyethylene vial (volume of vial is 20 ml).

In the case where validated method is applied [4], preliminary distillated samples must be enriched. Enrichment unit contains electrolytic cells, constant current supply unit and cooling unit. In one electrolysis run, a set of 16 electrolytic cells is connected in series and connected to a direct current source. Each enrichment run contains 16 samples, including 1 spike water (water of known tritium activity concentration used for enrichment factor determination). Each cell contains 250 ml of initial volume of the distillated sample. As an electrolyte, Na_2O_2 is used. In order to achieve approximately 10-15 ml of final volume of the samples, system works on 5 A. The enrichment operated with 5 A current could enrich approximately 1.7 g of water samples per hour. This means 5 days is necessary to enrich from 250 ml to 10-15 ml of water samples according to Faraday's factor.

An enrichment samples must be distilled again to remove electrolyte. After the second distillation an aliquot of 8 ml of the samples mix with 12 ml scintillation cocktail ULTIMA GOLD LLT in polyethylene vials.

For both methods, standard and validated, ultra-low-level liquid scintillation counter Quantulus 1220 is used. Measuring time of each sample is 300 min. Beside the samples, *dead water DW* (tritium free water) is also measured for background, as well as water of known tritium activity (*BEFORE BE*). *BEFORE*, after enrichment calls *SPIKE SW* water. Their ratio determines cell enrichment factor.

Counting efficiency is determined according to standard method [3]. For determination of efficiency, standard tritium solution ^{3}H 9031-OL-548/13 Czech Metrology Institute Type: ERX with activity 5.060 MBq on day 1.10.2013 is used. The tritium measurement uncertainty is expressed as the expanded measurement uncertainty for the factor k=2, which corresponds to a normal distribution with a confidence level of 95 %.

Minimum detectable activity (*MDA*) is measured using the following equation:

$$MDA = \frac{2{,}71 + 4{,}65\sqrt{R_b t_b}}{60 \varepsilon V t_b} \quad (1)$$

where R_b is background count rate (cps), t_b is background counting time (min), ε is efficiency (%) and V is volume of the *DW* (l). In case with enrichment, above equation is further divided by the electrolysis enrichment factor *Z*.

III. Section Title (e.g. Main Results)

For 15 natural water samples, validated method with electrolytic enrichment was performed in order to determined tritium activity. The obtained results are presented in Table 1. 1 drinking water, 9 precipitation samples and 5 surface water samples were analyzed. For the results shown from one electrolysis, it took 695 Ah to reduce initial volume of the samples to 13 ml (5.8 days at 5A). Calculated enrichment factor, *Z*, for this electrolysis is 13.2. Calculated *MDA* for this electrolysis is 0,23 Bq/l for background counts of 0.035 cps. Efficiency obtained by calibration of the spectrometer is 27.9 %.

Results for tritium activity in drinking water and precipitation from the reference meteorological station Zeleno Brdo are similar regardless of the month of sampling. Results obtained for precipitation at two locations in Vinča Institute of Nuclear Sciences, CS and MS are higher than results obtained for precipitation at ZB. This is in accordance with earlier results for tritium activity in precipitation for these locations [5-7]. In surface water samples, obtained results are different in relation to precipitation. Results for Sava and Danube are similar with values obtained for precipitation at MS. On the other hand, results for Mlaka Creek at location in Vinča Institute of Nuclear Sciences present higher values than surface water in Belgrade outside the Institute. These results are in accordance with previously results [5-7].

TABLE I

TRITIUM ACTIVITY CONCENTRATIONS IN WATER SAMPLES

Sample	Location	^{3}H (Bq/l) with enrichment	^{3}H (Bq/l) without enrichment
Drinking water	Belgrade	0.7 ± 0.2	
Precipitation	ZB IX 2020	1.2 ± 0.2	
	ZB X 2020	0.8 ± 0.2	5.4 ± 1.9
	ZB XI 2020	0.8 ± 0.2	
	CS IX 2020	3.1 ± 0.3	
	CS X 2020	3.0 ± 0.3	
	CS XI 2020	3.0 ± 0.3	
	MS IX 2020	1.6 ± 0.2	
	MS X 2020	1.7 ± 0.2	
	MS XI 2020	1.7 ± 0.2	
Surface water	Sava IX 2020	1.4 ± 0.2	
	Sava X 2020	1.2 ± 0.2	5.0 ± 1.9
	Danube X 2020	1.7 ± 0.2	5.0 ± 1.9
	Mlaka X 2020	5.4 ± 0.5	9.5 ± 2.0
	Mlaka XI 2020	5.4 ± 0.5	

For 4 samples: one precipitation, ZB, X 2020, and 3 surface water, Sava, X 2020, Danube, X 2020 and Mlaka Creek, X 2020 direct method for tritium analysis is performed. Results are presented in Table 1. As can be seen from the Table, values obtained for these samples which were measured directly are higher than values obtained for the same samples which were measured after enrichment. For this method, calculated *MDA* is 2.9 Bq/l for background counts of 0.033 cps. Efficiency is the same, 27.9 %.

For low level tritium activity it is evident that it is better to apply validated method, which also reduces the detection limit. In case it is certain that the activity of tritium is high the standard direct method can be applied. Especially because it is necessary to take into account that in case of expected high tritium values, contamination of the electrolysis system may occur.

In Serbia, legislation defines permitted values for tritum activity only for drinking water (100 Bq/l) [8]. In relation to this allowed value, all analyzed samples meet the criteria given by rulebook.

IV. Conclusion

Comparing standard and validated method for tritium determination in natural water samples, it can be concluded that validated method with electrolytic enrichment is more applicable for samples with low level tritium concentrations. Standard – direct method gives higher values and can be acceptable for expected high values of tritium. Method with enrichment also reduced minimum detectable activity. Liquid scintillation counting has proven to be irreplaceable in environmental tritium monitoring and for analysis of low level activity.

Acknowledgment

The research was funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

References

[1] "Investigation of the Environmental Fate of Tritium in the Atmosphere," in Part of the Tritium Studies Project, Info 0792, Canada, Minister of Public Works and Government Services Canada, Canadian Nuclear Safety Commission, 2009, Canada nuclear regulator, ISBN 978-1-100-13928-9

[2] Singh, A. Hussain, S.D. Khobragade "Analysis of isotope element by electrolytic enrichment method for ground water and surface water in Saurashtra region, Gujarat, India" Cogent Engineering, vol. 3, no. 1, 2016. DOI: 10.1080/23311916.2016.1164789

[3] ASTM D 4107-08 Standard Test Method for Tritium in Drinking Water

[4] Tecnical Procedure Note 19, Isotope Hydrology Lab., Procedure and Technique Critique for Tritium Enricment by Electrolysis at the IAEA Laboratory, IAEA 1976.

[5] M.M.Janković, D.J. Todorović, Z.Keleman, N.R. Miljević, "The measurement of tritium in water samples with electrolytic enrichment using liquid scintillation counter" Nuclear Technology & Radiation Protection, 27 (3) (2012), 239-246

[6] M.M. Janković, B.Ž. Janković, D.J. Todorović, Lj.M. Ignjatović, "Tritium concentration analysis in atmospheric precipitation in Serbia" Journal of Environmental Science and Health, Part A, 47(2012) 669-674.

[7] M.M. Janković, D.J. Todorović, "Examining the Relationships Between the Activities of ^{3}H in Precipitationand ^{137}Cs in Ground-Level Air in Belgrade City Area" Water Air & Soil Pollution, 223 (7) (2012) 4471-4483.

[8] Rulebook on Limits of Radionuclides Contentin Drinking Water, Foodstuffs, Feeding Stufffs, Medicines, Product for General Use, Construction Materials and Other Goods to be placed on Market (Official Gazette RS 36/18

HPGe detector efficiency optimization for the atypical measurement geometry of simulated aerosol filters

Jelena Krneta Nikolić*, Ivana Vukanac, Milica Rajačić, Dragana Todorović, Gordana Pantelić and Marija Janković

Abstract — **Gamma spectrometry is widely used method of choice for measurement of environmental samples conducted during monitoring of the environment and contamination control, as well as measurement of radionuclide content in various materials. However, one of the main challenges in this method of spectrometry is the determination of detection efficiency for different energies, different source-detector geometries and different composition of samples. This task is defined as an efficiency calibration of the detector. When using a commercial calibration sources is not possible, or the available sources are not adequate, the optimization of the efficiency calibration has to be performed.**

In this paper, the results of the optimization of efficiency calibration for the atypical geometry and composition of the simulated aerosol samples, measured within the Proficiency tests organized by International Atomic Energy Agency (IAEA), performed using EFFTRAN efficiency transfer software, will be presented and discussed.

Index Terms— **gamma spectrometry; efficiency calibration, EFFTRAN; optimization**

I. Introduction

Gamma spectrometry is one of the mostly often used measurement methods for determining the radionuclide content in various samples. It is a non-destructive method which can be applied for a wide range of environmental samples measured within the framework of a monitoring, as well as for contamination control.

This method is based on the interaction of gamma rays emitted from the sample with the active volume of the detector. Semiconductor detectors and among them, high purity germanium (HPGe) detectors, are mostly used due to their high sensitivity and good energy and time resolution. The result of a gamma spectrometric measurement is represented by the spectrum of photons originating from the source, that are collected by the multichannel analyzer (MCA) and the number of photons detected is proportional to the activity of the given radionuclide. The main challenge in this method of measurement is the determination of the detection efficiency which is dependent not only on the type of the detector and the energy of the emitted gamma photons, but also on the type of the measured sample: its geometry and chemical composition, the sample – detector geometry and the presence of the absorber. This task is defined as an efficiency calibration of the detector [1].

The most often used approach to the efficiency calibration is a direct measurement of different calibration sources containing γ-ray emitters and subsequent fitting of obtained results to a parametric function, thus obtaining the efficiency curve – a functional dependence of the efficiency with respect to the energy. Different sample types require different calibration curves. Due to that, this approach requires a large number of calibration sources, produced to mimic the real measured samples to the largest possible degree, which may not be available. This problem is especially pronounced when environmental samples are of interest due to their diversity in composition and structure [2].

When the sample of the atypical geometry or composition is presented, an optimization of the calibration curve using the means available in the laboratory has to be performed. One of the methods of optimization is the efficiency transfer using some type of software. The software calculates the efficiency transfer factors with which the original efficiency for a given energy needs to be multiplied in order to obtain the efficiency in the special case of the measured sample [3]. One of these software is EFFTRAN [4], a user friendly software that calculates efficiency transfer factors and coincidence summing correction factors for cylindrical samples.

In this paper, the results of the efficiency calibration for the atypical geometry and composition of the simulated aerosol samples, measured within the Proficiency tests organized by International Atomic Energy Agency (IAEA), performed using EFFTRAN efficiency transfer software,

Jelena Krneta Nikolić is with the University of Belgrade, Institute for Nuclear Sciences Vinča, National Institute of Republic of Serbia, Radiation and Environmental Protection Department Mike Petrovića Alasa 12-14, Vinča 11000 Belgrade, Serbia (e-mail: jnikolic@ vin.bg.ac.rs).

Ivana Vukanac is with the University of Belgrade, Institute for Nuclear Sciences Vinča, National Institute of Republic of Serbia, Radiation and Environmental Protection Department Mike Petrovića Alasa 12-14, Vinča 11000 Belgrade, Serbia (e-mail: vukanac@ vin.bg.ac.rs)

Milica Rajačić is with the University of Belgrade, Institute for Nuclear Sciences Vinča, National Institute of Republic of Serbia, Radiation and Environmental Protection Department Mike Petrovića Alasa 12-14, Vinča 11000 Belgrade, Serbia (e-mail: milica100@ vin.bg.ac.rs)

Dragana Todorović is with the University of Belgrade, Institute for Nuclear Sciences Vinča, National Institute of Republic of Serbia, Radiation and Environmental Protection Department Mike Petrovića Alasa 12-14, Vinča 11000 Belgrade, Serbia (e-mail: beba@ vinca.bg.ac.rs)

Gordana Pantelić is with the University of Belgrade, Institute for Nuclear Sciences Vinča, National Institute of Republic of Serbia, Radiation and Environmental Protection Department Mike Petrovića Alasa 12-14, Vinča 11000 Belgrade, Serbia (e-mail: pantelic@ vin.bg.ac.rs)

Marija Janković is with the University of Belgrade, Institute for Nuclear Sciences Vinča, National Institute of Republic of Serbia, Radiation and Environmental Protection Department Mike Petrovića Alasa 12-14, Vinča 11000 Belgrade, Serbia (e-mail: marijam@ vin.bg.ac.rs)

will be presented and discussed.

II. Main Results and Discussion

A. *Efficiency transfer*

The calculations of the efficiency transfer factors are based on the assumption that the detector efficiency for the special case of measured sample can be obtained by multiplying the reference efficiency (obtained by measuring the commercial or laboratory calibration source) by the efficiency transfer factors. In order to calculate these factors, a set of partial differential equations needs to be solved. For the purpose of the efficiency transfer, in our laboratory, EFFTRAN software is often used. It is organized as an user friendly Excel file with three modules. The software performs the needed calculations using a Monte Carlo integration, given the specific data are provided.

The data that the software requires are the detector characteristics (crystal material, diameter and length, thickness of the dead layer, housing geometry and composition, material of the window and window to crystal gap,) and the characteristics of both calibration sample used for the reference calibration curve and the measured sample (the diameter, filling height and thickness of the container, chemical composition and density of the sample matrix). . Because the model of the sample, as well as the detector crystal, can be constructed from cylinders only, the only complex operation required in the code is the calculation of the path length traversed through a cylinder of given dimensions by a gamma photon originating from an arbitrary location [4].

The choice of the reference efficiency plays a significant role in the final result, therefore it has to be chosen with care. This is especially important when the geometry of the measured sample differs significantly from the calibration source used for the efficiency calibration of the detector. Also the definition of the calibration source as well as the measured sample has to be performed as precise as possible, especially the chemical composition which has the largest influence on transfer factors. The final result of the calculation is the efficiency for the measured sample which is dependent on the reference efficiency used. The measurement uncertainty of the calculated efficiency is determined according to the following equation:
[5]:

$$u(\varepsilon) = \sqrt{(u(\varepsilon_{ref}))^2 + (u(C))^2 + (u_D)^2 + (u_S)^2} \quad (1)$$

where $u(\varepsilon)$ represents the combined measurement uncertainty of the efficiency for the measured sample, $u(\varepsilon_{ref})$ is the relative uncertainty of the reference efficiency value which has to be calculated, $u(C)$ is the uncertainty of the transfer factors calculated by the program as a statistical uncertainty of the Monte Carlo integration (≈ 1.2%), u_D is the uncertainty associated with the geometry of the detector and u_S is the uncertainty associated with the characteristics of the sample. The, for the measurement uncertainty of the measured activity, this component is combined with other contributions to obtain the total combined measurement uncertainty.

B. *Results and Discussion*

As it was said in the previous section, the final efficiency for the measured sample is dependent on the reference efficiency used. It is therefore crucial to perform some sort of validation of the results, when a choice of different reference efficiency is available.

In this investigation, the efficiency for an atypical geometry and composition has been calculated using three different reference efficiencies. The measured samples were simulated aerosol filters containing different artificial radionuclides, printed on a cellulose filter paper, diameter 43 mm, thickness of 1mm. These samples were measured within the World-Wide Open Proficiency test IAEA-TEL-2019-03, World-Wide Open Proficiency test IAEA-TEL-2020-03 and World-Wide Open Proficiency test IAEA-TEL-2020-05, organized by International Atomic Energy Agency (IAEA) during the year 2019 and 2020 [https://nucleus.iaea.org/sites/ReferenceMaterials/Pages/Interlaboratory-Studies.aspx]. The simulated aerosol filters contained Cs-134, Cs-134 (IAEA-TEL-2019-03 and IAEA-TEL-2020-05) and Ag-110m and Se-75 (IAEA-TEL-2020-03).

Three existing efficiency calibration curves were used for the reference efficiency: spiked charcoal in cylindrical geometry of 100 ml filled to a full, spiked mineralized grass in cylindrical geometry of 100 ml filled with 6.03g of matrix, and 50 ml vial, filled with 4.22g of aerosol [6]. The charcoal efficiency curve was used as it has the similar composition and density, the grass had the closest measurement geometry and the aerosol was used because it is readily used for the measurement of the prepared aerosol filters in the laboratory.

The simulated aerosol filters were measured on 2 p-type HPGe detectors. The duration of the measurement was 5100 s, 60000 s and 240000 s for the filter from IAEA-TEL-2019-03, IAEA-TEL-2020-03 and IAEA-TEL-2020-05 respectively. After the measurement, the activity of the present radionuclides was calculated using the grass matrix reference efficiencies (as it was the closest with the respect to the measurement geometry) in order to obtain the uncorrected results. Then the efficiency transfer was performed using EFFTRAN and the calculated transfer factors were applied in order to obtain the corrected result. Both uncorrected and corrected results were compared to the target value provided by the IAEA in the final report of the said Proficiency tests.

The uncorrected results, the corrected results and the target value for one simulated aerosol filter from each Proficiency test are presented in the Table I

TABLE I
THE RESULTS OF THE SIMULATED AEROSOL FILTER MEASUREMENTS USING DIFFERENT EFFICIENCIES AND THE TARGET VALUE, THE RESULTS WERE GIVEN WITH THE APPROPRIATE MEASUREMENT UNCERTAINTY, COVERAGE FACTOR 1

IAEA-TEL-2019-03					
Element	**Uncorrected result [Bq/sample]**	**Charcoal to filter efficiency transfer [Bq/sample]**	**Grass to filter efficiency transfer [Bq/sample]**	**Vial to filter efficiency transfer [Bq/sample]**	**Target value [Bq/sample]**
Cs-137	17.8 ± 0.7	9.6 ± 0.5	12.7 ± 0.6	13.1 ± 0.6	**13.02 ± 0.40**
Cs-134	21 ± 2	15 ± 1	20 ± 2	21 ± 2	**20.28 ± 0.61**
IAEA-TEL-2020-03					
Se-75	51 ± 2	23.5 ± 1.1	31 ± 2	29 ± 1	**31.3 ± 1.5**
Ag-110m	57 ± 2	30 ± 2	35 ± 2	35 ± 2	**35.1 ± 3.0**
IAEA-TEL-2020-05					
Cs-137	47.7 ± 0.8	25 ± 1	31 ± 1	29 ± 1	**28.6 ± 1.5**
Cs-134	27 ± 1	16.1 ± 0.7	20.4 ± 0.9	19.1 ± 0.9	**20.5 ± 1.1**

As it can be seen from the Table I, the uncorrected results differ significantly from the ones obtained using the efficiency transfer, although the composition of the mineralized grass (mainly cellulose and carbon) and the geometry were similar. Also, the transfer from the reference efficiency with the coal matrix produced the results that are significantly lower than the target value, meaning that the obtained efficiency is significantly overestimated. This can be explained by the large difference between the geometry of the reference efficiency which has greater diameter and sample height and therefore is the most diverse from the measured sample. Contrary to that, the transfer from the other two reference efficiency curves produced the results that are in agreement with the target values. For the elements that have multiple gamma lines, the coincidence correction factors were obtained using also EFFTRAN software. As it can be seen, the values for Cs-134, Se-75 and Ag-110m which are corrected for the coincidence summing effect and efficiency transfer from the grass reference efficiency proved to be the closest to the target value. For Cs-137, which has only one gamma emission and do not require coincidence summing correction, better results are obtained by transferring the aerosol reference efficiency. There is a local minimum at the energy of 661 keV in all efficiency curves regardless of the matrix of the calibration source. This leads to underestimation of the efficiency for this energy, which in turn produces an underestimated transferred efficiency. The recommendation for this energy is to use the efficiency obtained directly from the calibration source measurement, rather than from the calibration curve. Also, it is evident that the aerosol calibration source, although it closely represents the real aerosol samples, is not the best choice for the simulated aerosol filters which have an atypical geometry and composition. The mineralized grass calibration source proves to be the best reference calibration for the efficiency transfer since its diameter is very close to the diameter of the measured sample and more important, its thickness and chemical composition are virtually the same.

All the results obtained by using the efficiency transfer from the grass and aerosol matrix are acceptable, while none of the uncorrected results are acceptable. This obviously proves that the efficiency transfer has to be performed with the adequate reference calibration curve.

III. CONCLUSION

In this paper we presented the optimization of the efficiency calibration of HPGe detectors for the measurement of the simulated aerosol filters, measured within three Proficiency tests organized by IAEA. In case of the atypical geometry and composition of the measured sample, the efficiency transfer is inevitable, since the uncorrected activities are not in agreement with the target values, although the calibration source used for the efficiency calibration is of the similar geometry and composition. The choice of the reference efficiency curve for the efficiency transfer should be based on the similarities between the thickness and composition of the calibration source and the measured sample, since this choice produces the best results.

ACKNOWLEDGMENT

The research was funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

REFERENCES

[1] G. Gilmore, "*Practical Gamma-ray Spectrometry, Second edition,* New York, USA Wiley, 2008,

[2] M.J. Daza, B. Quintana, M. Garcia Tavalera, F. Fernandez, *"Efficiency calibration of a HPGe detector in the 46.54-2000 keV energy range for the measurement of environmental samples"*, Nucl. Instr. Method. A, vol. 470, pp. 520–532, 2001

[3] L. Moens, J. De Donder, X. Lin et al., *"Calculation of the absolute peak efficiency of gamma-ray detectors for different counting geometries"*, Nucl. Instr. Method, vol. 187, pp. 451-472, 1981

[4] T. Vidmar, *"EFFTRAN – A Monte Carlo Efficiency Transfer Code For gamma-ray spectrometry"*, Nucl. Instr. Method. A, vol. 550,pp. 603-608, 2005

[5] *IAEA-TECDOC-1401*, "*Quantifying uncertainty in nuclear analytical measurements"*, Vienna, Austria, IAEA, 2004

[6] J. Krneta Nikolic, M. Rajacic, D. Todorovic, M. Jankovic, N. Sarap, G. Pantelic, I. Vukanac, *"Semiempirical Efficiency Calibration in Semiconductor HPGe Gamma-Ray Spectroscopy"*, Journal of

Spectroscopy, vol 2018, Article ID 5392658, 2018

РАЧУНАРСКА ТЕХНИКА И ИНФОРМАТИКА
/
COMPUTING AND INFORMATION ENGINEERING

(РТ/RTI)

Jedno rješenje posrednika u sistemu uslovnog pristupa digitalne televizije

Radenko Banović, Ilija Bašičević i Nemanja Lazukić

***Apstrakt*— Postoje dvije vrste TV (televizijske) usluge u domenu pretplate: javna (svima dostupan sadržaj za koji nije potrebna pretplata) i pretplatnička TV usluga (TV sadržaj je dostupan samo pretplaćenim korisnicima). Da bi pretplatnička TV usluga imala smisla potrebno je zaštititi TV sadržaj cijelim prenosnim putem. Postoji nekoliko modela zaštite pretplatničkog TV sadržaja, a jedan od njih je CAS (eng. Conditional Access System). Kompanija Widevine je kreirala rješenje sistema uslovnog pristupa (CAS) takvo da je besplatno za sve operatere. Da bi operateri mogli upravljati korisnicima i sadržajem, potrebno je implementirati korisničku upravljačku logiku sistema. U ovom radu je predstavljeno jedno rješenje softverskog posrednika (eng. Proxy) u kome je realizovana korisnička upravljačka logika sistema uslovnog pristupa u Widevine CAS sistemu.**

***Ključne reči*—Conditional Access System; Proxy; Digital Television;**

I. UVOD

Pretplatnička televizija je usluga koju nude satelitski, kablovski i drugi distributeri televizijskih kanala. Ključna tačka preduzetničkog modela u pretplatničkoj televiziji jeste zaštita televizijskog sadržaja cijelim prenosnim putem, od emitera do krajnjeg korisnika, čime se otklanja mogućnost pristupa sadržaju nepretplaćenim korisnicima[1]. Postoji nekoliko tehnologija zaštite televizijskog sadržaja, a najpoznatije su upravljanje digitalnim pravima (eng. Digital Rights Management) i sistem uslovnog pristupa (eng. Conditional Access System).

Sistem uslovnog pristupa predstavlja zaštitu prenosnog puta[2] (i on se najčešće koristi za televiziju uživo), dok je upravljanje digitalnim pravima zamišljeno kao mehanizam zaštite sadržaja (te se najčešće koristi za televiziju na zahtjev (eng. On Demand)). Za razliku od upravljanja digitalnim pravima, u sistemima uslovnog pristupa uobičajno je da se nakon određenog vremenskog intervala mijenjaju ključevi kojima je skremblovan sadržaj koji se dostavlja korisniku[3]. Kompanija Widevine je kreirala sopstveno rješenje CAS sistema za Android TV i Android STB (Set-Top Box) uređaje koje je napravilo veliki pomak u industriji digitalne televizije.

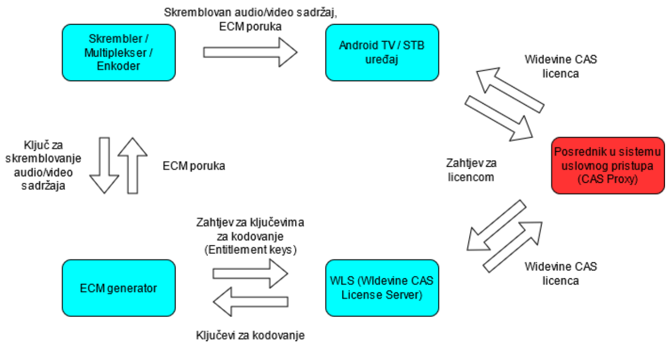
Radenko Banović – Fakultet Tehničkih Nauka, Univerzitet u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: Radenko.Banovic@rt-rk.com).

Ilija Bašičević – Fakultet Tehničkih Nauka, Univerzitet u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: ilibas@uns.ac.rs).

Nemanja Lazukić – Istraživačko-razvojni Institut RT-RK, Novi Sad, Srbija, (e-mail: Nemanja.Lazukic@rt-rk.com).

II. WIDEVINE CAS SISTEM

Ključna prednost Widevine CAS rješenja u odnosu na konkurenciju jeste to što je kompletan CAS ekosistem dat operatorima na besplatno korištenje, pod uslovom da se izvršava na Android TV operativnom sistemu[5]. Komponente Widevine CAS sistema su : licencni poslužioc (eng. License Server), OEMCrypto (modul koji se integriše u Android TV), ECM (eng. Entitlement Control Message) generator, skrembler i posrednik u sistemu uslovnog pristupa. Sl. 1. prikazuje komponente Widevine CAS sistema na visokom nivou apstrakcije.



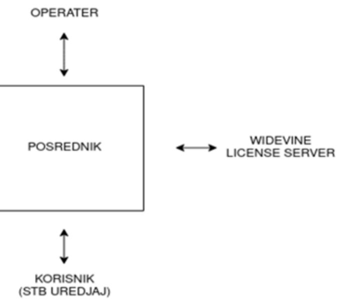
Sl. 1. Widevine CAS sistem

Neke komponente sistema su date tako da se ne mogu prilagođavati (OEMCrypto, licencni poslužioc), dok ECM generator i posrednik u sistemu uslovnog pristupa moraju da se implementiraju za svakog provajdera posebno, ali tako da se oslanjaju na Widevine SDK (eng. Software Development Kit). Licencni poslužioc je ključna tačka sistema u kojoj se sastaju predajna i prijemna strana. Korištenje licencnog poslužioca je moguće nakon što Widevine odobri zahtjev za korištenjem, i kreira posebne URL putanje prema zahtjevu provajdera.

Sa predajne strane se licencnom poslužiocu šalje zahtjev za dostavljanjem ključa (eng. Entitlement Key) kojim se enkriptuje ECM poruka u kojoj se nalaze ključevi kojima je skremblovan televizijski sadržaj. Sa prijemne strane se licencnom poslužiocu šalje zahtjev za dostavljanjem licence iz koje se izvlače ključevi kojima je moguće dekriptovati ECM poruku, te sa ključevima izvučenim iz ECM poruke deskremblovati televizijski sadržaj i prikazati ga korisniku.

III. POSREDNIK U SISTEMU USLOVNOG PRISTUPA SA KORISNIČKOM UPRAVLJAČKOM LOGIKOM

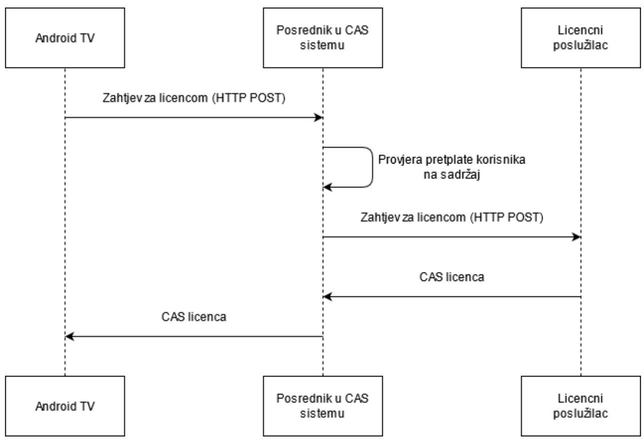
Posrednik kao jedan od elemenata CAS sistema komunicira sa Android TV / STB uređajem, kao i sa licencnim poslužiocem. Takođe, potrebno je kreirati korisnički interfejs preko kog je moguće unositi podatke vezane za korisnike, kanale, pakete na koje su korisnici pretplaćeni, što je ilustrovano u Sl. 2.



Sl 2. Interakcija posrednika sa okolinom

Posrednik je zamišljen kao mrežno orijentisan servis koji koristi REST (eng. Representational State Transfer) API (eng. Application Programming Interface) arhitekturu softvera[4]. Korištenje REST API arhitekture je omogućilo identifikovanje različitih resursa uz pomoć definisanja posebnih URI (eng. Uniform Resource Identifier) putanja, tako da i operater i korisnik (STB uređaj) mogu koristiti posrednik šaljući različite zahtjeve ka njemu. URI putanje namijenjene komunikaciji sa operaterom se odnose na upravljanje sadržajem baze podataka (dodavanje i brisanje korisnika, uređaja, paketa kanala, ažuriranje informacija o pretplatama korisnika).

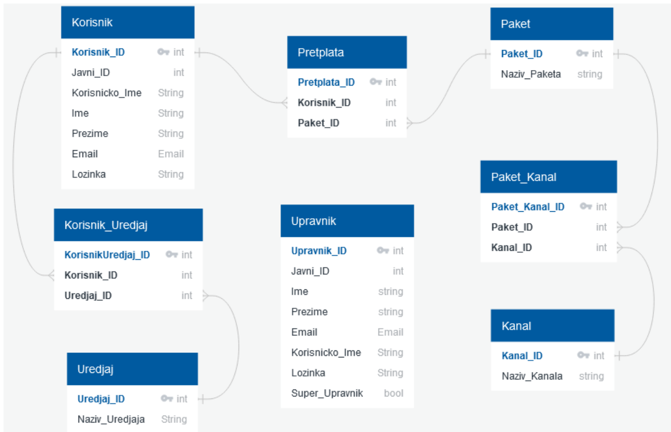
STB uređaj korisniku šalje zahtjev za licencom, zatim se nakon obrade zahtjeva provjerava da li je korisnik koji zahtjeva licencu za svoj STB uređaj pretplaćen na željeni sadržaj. Ukoliko jeste pretplaćen, zahtjev se prosljeđuje licencnom serveru, te se licenca dobijena od strane licencnog servera prosljeđuje korisniku koji je uputio zahtjev za licencom. Komunikacija između STB uređaja i posrednika, te posrednika i licencnog poslužioca je prikazana u Sl. 3



Sl. 4. Dijagram poziva posrednika

IV. OPIS REALIZACIJE

Posrednik je ralizovan kao HTTP poslužilac u C++ programskom jeziku, jer je SDK na koji se on oslanja takođe realizovan u C++ programskom jeziku. Jezgro posrednika predstavlja baza podataka u kojoj se nalaze sve relevantne informacije na osnovu kojih je moguće odrediti da li je korisnik pretplaćen na odgovarajuće pakete kanala. Šema baze podataka je prikazana na Sl. 4.



Sl. 4. Šema baze podataka

A. *Alati korišteni za realizaciju*

Pošto je C++ izabran kao programski jezik, a posrednik treba da bude HTTP poslužilac izabrali smo Mongoose biblioteku[6] u kojoj je implementiran na događaj pobuđeni (eng. Event-driven) neblokirajući API za HTTP i uz koji je moguće kreirati REST API servise koji su neophodni za komunikaciju sa okolinom. Za upravljanje bazom podataka korištena je SQlite biblioteka[7].

B. *Implementacija obrade zahtjeva za licencom*

Nakon što posrednik zaprimi zahtjev na URI putanji *dobavi_licencu* u funkciji *handle_lic_req()* se uz pomoć poziva SDK funkcije *getDeviceInfo()* iz zahtjeva za licencu dobijaju informacije o uređaju i to : proizvođač, model, identifikacioni broj uređaja i serijski broj sertifikata uređaja. Iz zahtjeva za licencu se uz pomoć poziva SDK funkcije *getContentId()* dobavlja informacija o paketu kanala za koji se šalje zahtjev za licencu.

Na osnovu dobijenih informacija o uređaju iz baze podataka se dobavlja informacija o korisniku. Zatim se na osnovu informacije o korisniku i paketu kanala provjerava da li je korisnik pretplaćen na željeni paket kanala. Ukoliko je korisnik pretplaćen na paket kanala pozivom SDK funkcije *GenerateLicenseRequestAsJSON()* se na osnovu zahtjeva za licencu generiše zahtjev koji se preko HTTP Post metode korištenjem Curl biblioteke šalje licencnom serveru.

HTTP odgovor dobijen on licencnog poslužioca se prosljeđuje STB uređaju koji je poslao zahtjev pozivom funkcije mg_printf() biblioteke mongoose. Ukoliko korisnik nije pretplaćen na željeni sadržaj na STB uređaj se odmah šalje HTTP odgovor sa statusnim kodom 405 koji se odnosi na to da takav zahtjev nije dozvoljen.

Svaka akcija za popunjavanje baze podataka je kreirana sa posebnom uri putanjom i funkcijom koja obrađuje zahtjev. Akcije koje su obrađene su : dodaj korisnika, obriši korisnika, pretplati korisnika, ukini pretplatu korisnika, dodaj uređaj, obriši uređaj, dodaj kanal, obriši kanal, dodaj pretplatu, obriši pretplatu, dodaj kanal u paket i izbaci kanal iz paketa. U ovoj fazi razvoja nije predviđena realizacija prednjeg dijela (eng. Front-end) zbog čega su implementirane samo funkcije za popunjavanje baze podataka, i čitanja iz baze podataka neophodna za dobavljanje licence.

V. TESTIRANJE

Predloženo rješenje je testirano na Synaptics BG5CT STB (Sl. 8.) uređajima sa operativnim sistemom Android Q. Korištena je Live Channels korisnička aplikacija koja se oslanja na Comedia DTV (eng. Digital Television) srednji sloj kompanije iWedia, u kom je integrisam OEMCrypto koji kreira zahtjev za licencom i koji služi za deskremblovanje televizijskog sadržaja.

Sl. 8. Synaptics BG5CT platforma

Sa predajne strane je korišten TSDuck set alata [8] koji se u ovom slucaju koristio za skremblovanje TS (eng. Transport Stream) toka podataka koji se nalazio u izvorišnoj datoteci, kao i za slanje skremblovanog toka podataka ka odredničnom STB uređaju korištenjem računarske mrežne infrastrukture i IPv6 (eng. Internet Protocol version 6) protokola. Takođe, sa predajne strane je korišten ECM generator koji je razvijan u paraleli sa posrednikom u sistemu uslovnog pristupa.

Funkcionalnost je testirana korištenjem više prijemnih uređaja pri čemu su mijenjane informacije o pretplaćenim korisnicima u bazi podataka. Kreirano je nekoliko testnih slučajeva u kojima su različiti korisnici u različitim testnim slučajevima bili pretplaćeni na željeni sadržaj. Jedan primjer testnog slučaja: korisnik A je pretplaćen na željeni sadržaj, korisnik B nije pretplaćen na željeni sadržaj, testiranjem je utvrđeno da korisnik A ima pristup sadržaju, dok korisnik B nema pristpu sadržaju.

Po završetku testiranja utvrđeno je da su STB uređaji pretplaćenih korisnika uspješno deskremblovali i reprodukovali sadržaj iz izvorišne datoteke koja je emitovana ka njima. U slučajevima nepretplaćenih korisnika STB uređaji su dobijali odgovor od posrednika da nisu pretplaćeni na željeni sadržaj, te nisu dobili licencu iz koje bi mogli izvuću ključeve kojima bi uspješno deskremblovali sadržaj.
Testiranjem je utvrđena funkcionalnost rješenja.

VI. ZAKLJUČAK

U ovom radu je prikazano jedno rješenje posrednika u sistemu uslovnog pristupa sa korisničkom upravljačkom logikom. Opisan je Widevine CAS sistem u cjelini kao i uloga posrednika u njemu. Navedeni su svi alati korišteni u realizaciji rješenja, te je dat detaljan opis rješenja. Rješenje je testirano korištenjem nekoliko prijemnih uređaja i nakon uspješno završenih testova potvrđena je funkcionalnost rješenja. Doprinos ovog rada u odnosu na postojeća rješenja je u tome što je kompatibilan sa Widevine CAS ekosistemom. U budućnosti ovo rješenje može biti unaprijeđeno kreiranjem prednjeg dijela poslužioca čime bi se omogućio jednostavan vizuelni prikaz i lakše upravljanje pretplatom korisnika, te proširenjem zadnjeg dijela poslužioca.

LITERATURA

[1] I. Kaštelan, V. Peković, V. Zlokolica, J. Zloh, D. Trifunović, “Simultaneous automated verification of conditional access system on multiple TV sets,” Proc. IEEE International Conference on Consumer Electronics, Berlin, Germany, pp. 269-270, Sept. 2012.

[2] Fu-Kuan Tu, Chi-Sung Laih and Hsu-Hung Tung, "On key distribution management for conditional access system on pay-TV system," in *IEEE Transactions on Consumer Electronics*, vol. 45, no. 1, pp. 151-158, Feb. 1999, doi: 10.1109/30.754430.

[3] Milan Bjelica, Nikola Teslić, Velibor Mihić, “Softver u digitalnoj televiziji 1”, 2017.

[4] Fielding, Roy Thomas (2000). "Chapter 5: Representational State Transfer (REST)". Architectural Styles and the Design of Network-based Software Architectures (Ph.D.). University of California, Irvine

[5] Widevine CAS, Jun 2021. [online]. https://www.widevine.com/solutions/widevine-cas

[6] Mongoose - Embedded Web Server, Jun 2021. [online]. https://github.com/cesanta/mongoose

[7] SQLite, Jun 2021. [online]. https://www.sqlite.org

[8] TSDuck, Jun 2021. [online]. https://tsduck.io/

ABSTRACT

There are two types of TV (television) services in the domain of subscription: public (content available to all for which no subscription is required) and subscriber TV service (TV content is available only to subscribed users). In order for the subscriber TV service to make sense, it is necessary to protect the TV content throughout the transmission. There are several models of protection of subscriber TV content, and one of them is CAS (Conditional Access System). Widevine has created a conditional access system (CAS) solution that is free for all operators. In order for operators to be able to manage users and content, it is necessary to implement the user management logic of the system. This paper presents a solution of a proxy server in which the user control logic of the conditional access system in the Widevine CAS system is realized.

One solution of proxy server in the digital television conditional access system

Radenko Banović, Ilija Bašičević, Nemanja Lazukić

Jedno rješenje ECM generatora

Radenko Banović, Ilija Bašičević, Ksenija Popov i Milenko Maksić

***Apstrakt*—Zaštita televizijskog sadržaja predstavlja jedan od najvećih izazova u industriji digitalne televizije usljed sve manjeg broja televizijskih kanala čije se gledanje ne naplaćuje. Da bi omogućili naplaćivanje televizijskog sadržaja korisnicima, potrebno je zaštiti televizijski sadržaj cijelim prenosnim putem. Najkorišteniji model zaštite živog televizijskog sadržaja je CAS (eng. Conditional Access System). CAS model podrazumijeva postupak zaštite video i audio sadržaja skremblovanjem koje ima za cilj sprječavanje neovlaštene reporodukcije audio i video sadržaja. Kontrolne riječi kojima je izvršeno skremblovanje se prenose istim prenosnim kanalom kao i skremblovani sadržaj u okviru ECM (eng. Entitlement Control Message) poruke ali u enkriptovanom obliku. Kompanija Widevine je realizovala sopstveni CAS ekosistem potpuno besplatan za sve korisnike. U ovom radu je predstavljeno jedno rješenje ECM generatora u Widevine CAS sistemu.**

***Ključne reči*—ECM generator, Conditional Access System, Digital Television**

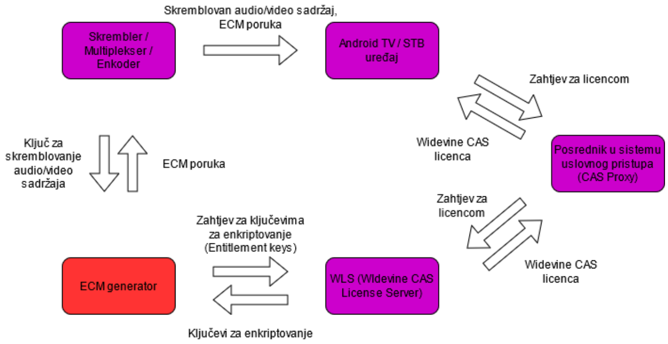
I. Uvod

U junu 2014. godine je prvi put prikazan Andoid TV operativni sistem, koji je prilagođena verzija Android operativnog sistema za televizore i STB (set-top box) uređaje[4]. Do danas je veliki broj proizvođača televizora i STB uređaja, kao i operatera televizijskih kanala integrisalo Android TV kao operativni sistem koji se izvršava na njihovim uređajima[5].

Kako bi privoljeli i preostale operatere televizijskih kanala i proizvođače televizora i STB uređaja da integrišu Android TV na svoje uređaje kreiran je Widevine CAS sistem uslovnog pristupa koji je na korištenje dat potpuno besplatno, ali može da se koristi samo uz Android TV operativni sistem. Da bi sistem postao funkcionalan, potrebno je implementirati ECM generator i posrednik u sistemu uslovnog pristupa, za šta je dat SDK (eng. Software Development Kit).

Komponente sistema koje je potrebno implementirati nisu implementirane da bi svaki operater televizijskih kanala prilagodio sistem svojim potrebama. Postoji nekoliko primjera implementacije ECM generatora [1, 2], ali oni nisu prilagođeni Widevine CAS ekosistemu. Widevine CAS sistem je prikazan u Sl. 1.

Radenko Banović – Fakultet Tehničkih Nauka, Univerzitet u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: Radenko.Banovic@rt-rk.com).
Ilija Bašičević – Fakultet Tehničkih Nauka, Univerzitet u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: ilibas@uns.ac.rs).
Ksenija Popov – Istraživačko-razvojni Institut RT-RK, Novi Sad, Srbija, (e-mail: Ksenija.Popov@rt-rk.com).
Milenko Maksić – Istraživačko-razvojni Institut RT-RK, Novi Sad, Srbija, (e-mail: Milenko.Maksic@rt-rk.com).



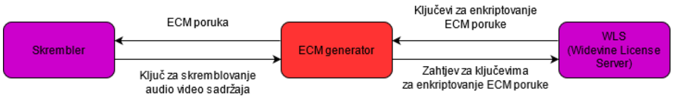
Sl. 1. Widevine CAS sistem

II. ECM generator

Da bi audio i video sadržaj bio zaštićen tokom kompletnog prenosnog toka vrši se postupak skremblovanja. Inverzni postupak u odnosu na skremblovanje se naziva deskremblovanje, njime se zaštićeni sadržaj prevodi u osnovni format razumljiv audio i video dekoderima[3].

Skremblovanje se vrši korištenjem kontrolne riječi (ključa za skremblovanje). Korištenje kontrolne riječi u procesu skremblovanja omogućuje promjenu kontrolne riječi u vremenu, a period između dve promjene se naziva periodom kriptovanja. Što je češća izmjena kontrolne riječi, to je proces skremblovanja bezbjedniji.

Trenutno korištena kontrolna riječ prenosi se u okviru ECM poruke koja se generiše u ECM generatoru. PID (eng. Packet Identifier) vrijednost TS (eng. Transport Stream) paketa u kom se nalazi ECM poruka se nalazi u CA deskriptoru PMT tabele. ECM generator u Widevine CAS sistemu komunicira sa licencnim poslužiocem (eng. License Server) i skremblerom. Pozicija ECM generatora u Widevine CAS sistem je prikazan na Sl. 2.

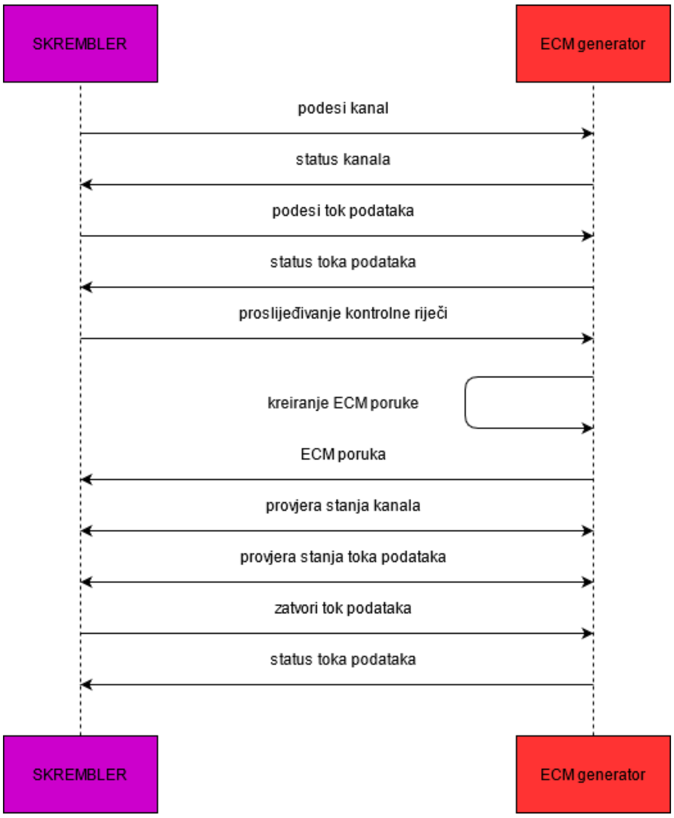


Sl. 2. Pozicija ECM generatora u widevine sistemu

U ovom radu je korišten skrembler implementiran u TS Duck programskoj podršci. TS Duck je set alata koji se koristi za manipulaciju MPEG prenosnim tokovima, a jedan od alata je i skrembler koji može da koristi i eksterni ECM generator za generisanje ECM poruka[6]. ECM generator i TS Duck komuniciraju po ECMG/SCS (eng. Simulcrypt Synchroniser) protokolu[7].

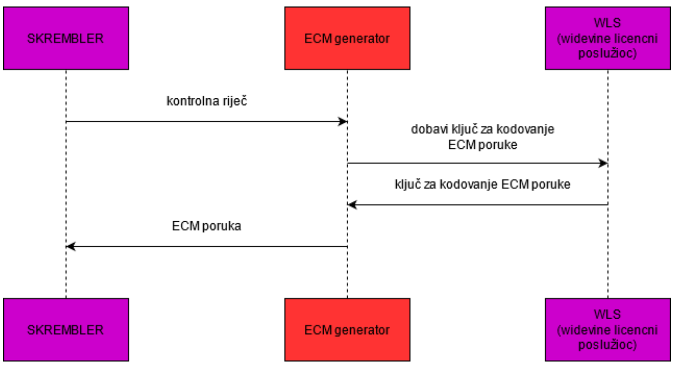
U komunikaciji između generatora i skremblera, skrembler je implementiran kao klijent, dok generator treba da bude implementiran kao poslužioc, te je ECM generator je u smislu komunikacije sa skremblerom implementiran kao TCP/IP poslužioc koja čeka zahtjeve od skremblera.

Po uspostavljanju veze generatora i skremblera kreira se sesija koja je zadužena za razmjenu poruka u okviru ECMG/SCS protokola. ECMG/SCS protokol je prikazan na Sl. 3.



Sl 3. Dijagram ECMG/SCS protokola

WLS (eng. Widevine License Server) je već gotovo rješenje sa kojim ECM generator komunicira putem HTTP protokola. ECM generator treba od WLS da dobavi (eng. Entitlement) ključeve kojima će enkriptovati ECM poruku u kojoj se nalaze ključevi kojima je skremblovan sadržaj, tako da u slučaju presretanja ECM poruke presretač ne može da dobije informaciju o ključevima kojima je sadržaj skremblovan. Dijagram komunikacije skremblera, ECM generatora i WLS je prikazan na Sl. 4.



Sl. 4. Dijagram komunikacije ECM generatora

III. Opis realizacije

ECM generator je realizovan kao C++ CLI (eng. Command Line Interface) aplikacija. Prilikom pokretanja aplikacije potrebno je proslijediti broj porta na kom aplikacija osluškuje zahtjev klijenta (skrembler) za uspostavljanjem veze. Kompletno rješenje se oslanja na Widevine SDK (eng. Software Development Kit). Rješenje možemo podijeliti u tri logičke cjeline, i to: TCP/IP poslužilac, ECMG/SCS protokol, i ECM generator.

A. TCP/IP poslužilac

Ovaj modul sadrži dvije funkcije*: void TCPstart(int port, void (*onSessionEstablished)())* i *void TCPstop(int port).* Funkcija *TCPstart* kreira TCP/IP utičnice sa podrškom za IPv4 i IPv6 protokole, stavlja poslužioca u stanje čekanja zahtjeva za konekcijom klijenta, te uspostavlja vezu i kreira sesiju za korisnik / poslužilac komunikaciju. Funkcijom *TCPstop* se zatvara otvorena sesija.

B. ECMG/SCS protokol

U ovom modulu je implementiran ECMG/SCS protokol. Implementiran je tako da se izvršava u *while* petlji, poziva se funkcija *read()* koja je blokirajuća, i koja zaustavlja izvršavanje petlje dok se memorija za smještanje dolaznih podataka ne popuni. Iz pristiglih podataka se pozivom funkcije *int32_t msg_pars(const uint8_t* buff, uint32_t size, struct ecmgp_msg* msg)* popunjava sktruktura *ecmgp_msg*.

Jedno od polja strukture koja predstavlja poruku je tip poruke, na osnovu kog se korištenjem *swich* grananja bira grana u kojoj se priprema odgovor na poslatu poruku. Tip poruke može biti : *CHANNEL_SETUP*, *STREAM_SETUP*, *CW_PROVISION*, *STREAM_CLOSE_REQUEST*. Svaka od grana popunjava strukturu koja predstavlja poruku, te se poziva funkcija *int32_t msg_generator(uint8_t* buff, struct ecmgp_msg* msg)* koja od podataka iz strukture kreira poruku koja se šalje ka klijentu.

Poruka tipa *CW_PROVISION* nosi i vrijednost ključa za skremblovanje audio/video sadržaja koja kriptovana treba da se nađe u ECM poruci. U funkciji *int32_t gen_ecm_datagram(uint8_t* ecm_datagram, struct ecmgp_msg* msg)* je implementirano kreiranje ECM poruke, te se ona poziva u grani obrade poruke tipa *CW_PROVISION.* Nakon poziva ove funkcije ECM poruka se dodaje kao polje strukture *ecmgp_msg*, poziva se funkcija *msg_generator* nakon koje se kreirana ECM poruka šalje skrembleru.

C. ECM generator (u užem smislu)

Za generisanje ECM poruke i kreiranje TS paketa, te kreiranje zahtjeva za ključevima za enkriptovanje ECM poruke i parsiranjem odgovora dobijenog od WLS korištene su funkcije dobijene iz Widevine SDK paketa. Funkcija u kojoj se kreira ECM poruka *gen_ecm_datagram()* kroz parametar dobija poruku dobijenu od skremblera u kojoj se nalaze ključevi za skremblovanje audio/video podataka. Pored ključeva za skremblovanje, potrebno je dobaviti ključeve za enkriptovanje ECM poruke koji se dobijaju slanjem ispravnog HTTP zahtjeva ka WLS.

Pozivom funkcije *CreateEntitlementRequest()* koja je dio Widevine SDK kreira se zahtjev za ključevima za enkriptovanje ECM poruke. Da bi se kreirao ispravan zahtjev potrebno je funkciji proslijediti sledeće podatke: identifikator sadržaja za skremblovanje, naziv operatera, broj ključeva za skremblovanje (jedan, ili dva), rezolucija sadržaja, ime operatera koji potpisuje zahtjev za licencom, ključ za potpisivanje enkriptovanog zahtjeva i vektor za potpisivanje zahtjeva. Nakon što je zahtjev ispravno kreiran korištena je CURL biblioteka[8] kako bi se poslao HTTP zahtjev ka WLS, nakon čega se odgovor upisuje u željeni dio memorije.

Nakon dobijenog odgovora, poziva se funkcija *ParseEntitlementResponse()* koja iz sirovog odgovora izvlači dva ključa za enkriptovanje ECM poruke. Pozivom funkcije *GenerateEcm()* kojoj se kao parametri proslijeđuju ključevi za enkriptovanje ECM poruke kreira se ECM poruka, da bi se na kraju pozivom *GenerateTsPacket()* kreirao paket koji se šalje ka skrembleru.

IV. Testiranje

U paraleli sa izradom ECM generatora, kreirano je i rješenje posrednika u sistemu uslovnog pristupa (CAS Proxy), te je integrisana Widevine OEMCrypto biblioteka u Android STB uređaj. Nakon što na Android STB uređaj pristigne skremblovan audio/video sadržaj, on posredniku u sistemu uslovnog pristupa šalje zahtjev za licencom, sa informacijom o kom sadržaju je riječ. Ukoliko dobije odgovor od posrednika, licenca se proslijeđuje OEMCrypto biblioteci koja iz licence izvlači ključeve za dekriptovanje ECM poruke. Ukoliko se poruka uspješno dekriptuje, ključevima dobijenim iz ECM poruke se deskrembluje audio/video sadržaj, te je na ekranu moguće vidjeti audio/video sadržaj koji je poslat na STB uređaj. Predloženo rješenje je testnirano na Synaptics BG5CT STB (Sl. 5.) uređajima sa operativnim sistemom Android Q. Korištena je Live Channels korisnička aplikacija koja se oslanja na Comedia DTV (eng. Digital Television) srednji sloj kompanije iWedia.

Sl. 5. Synaptics BG5CT platforma

Pošto su STB uređaji uspješno deskremblovali i reprodukovali audio/video sadržaj skremblovan ključevima generisanim u TS Duck alatu, te ECM porukama enkriptovanim ključevima dobijenim od WLS, konstatovali smo da je testiranjem utvrđena funkcionalnost rješenja.

V. Zaključak

U ovom radu je prikazano jedno rješene ECM generatora u Widevine CAS sistemu. U uvodu je objašnjena uloga i značaj CAS sistema, kao i njegova komercijalna primjena. Bliže je opisan način zaštite televizijskog sadržaja u CAS sistemima, Prikazan je opis rješenja. Rješenje je testirano na nekoliko prijemnih uređaja, sa nekoliko ulaznih tokova podataka te je potvrđena funkcionalnost sistema. U budućnosti se ovo rješenje može unaprijediti podrškom za kreiranje ECM poruke za više različitih tokova podataka u paraleli.

Literatura

[1] Li Xi and Chen Xin, "Design of Digital Video Broadcasting Conditional Access System Headend Communication Interface," in *Computer and Modernization*, vol. 1, no. 3, pp. 118-121, 2012.

[2] In-Hee Jo and Byoung-Soo Koh, " Building a common encryption scrambler to protect paid broadcast services," in *International Journal of Internet Technology and Secured Transactions*, vol. 6, no. 3, Nov. 2016, doi: 10.1504/IJITST.2016.080391.

[3] Milan Bjelica, Nikola Teslić, Velibor Mihić, "Softver u digitalnoj televiziji 1", 2017.

[4] Google Unveils First Android TV Device, Jun 2021. [online]. https://www.nexttv.com/news/google-unveils-first-android-tv-device-384772

[5] Android TV OS reaches 80M monthly active devices, Jun 2021. [online]. https://techcrunch.com/2021/05/18/android-tv-os-reaches-80m-monthly-active-devices-adds-new-features/

[6] TSDuck, Jun 2021. [online]. https://tsduck.io/

[7] ECMG/SCS protokol, Oktobar 2008. [online]. https://www.etsi.org/deliver/etsi_ts/103100_103199/103197/01.05.01_60/ts_103197v010501p.pdf

[8] Curl library, Jun 2021. [online]. https://curl.se/

Abstract

The protection of television content is one of the biggest challenges in the digital television industry due to the declining number of free-to-air television channels. In order to enable the charging of television content to operators, it is necessary to protect television content throughout the transmission. The most widely used model for the protection of live television content is the CAS (Conditional Access System). The CAS model involves a process of protecting video and audio content by scrambling that aims to prevent unauthorized reproduction of audio and video content. The scrambled control words are transmitted via the same transmission channel as the scrambled content within the ECM (Entitlement Control Message) message but in encrypted form. Widevine has implemented its own CAS model completely free for all users. In this paper, one solution of ECM generator in Widevine CAS system is presented.

One solution of ECM generator

Radenko Banovic, Ilija Basicevic, Ksenija Popov, Milenko Maksic

Aplikacija za demonstraciju *XSS* sigurnosnih propusta

Katarina Simić, Žarko Stanisavljević

***Apstrakt* — *XSS* (eng. *Cross-site scripting*) je jedna od najčešćih ranjivosti veb aplikacija uprkos tome što postoji veliki broj različitih mehanizama zaštite. U ovom radu prikazana je implementacija jedne ranjive aplikacije u okviru koje je moguće demonstrirati različite tipove XSS sigurnosnih propusta, kao i načina njihove zloupotrebe, ali i eliminisanja. Aplikacija se može koristiti kao edukativno sredstvo za praktičnu obuku softverskih inženjera u zatvorenom i bezbednom okruženju.**

***Ključne reči* — XSS, OWASP top 10, sigurnosni propusti.**

I. Uvod

Važnost veb aplikacija posebno je došla do izražaja u trenutku pandemije koronavirusa koja je počela 2019. godine i još uvek traje. Od tada ljudi sve više završavaju svoje poslove i obavljaju određene aktivnosti, poput online kupovine ili korišćenja društvenih mreža, uz pomoć veb aplikacija. Na ovaj način korisnici na različitim mestima ostavljaju svoje poverljive podatke, verujući veb aplikacijama da oni neće pasti u pogrešne ruke. Iz ovog razloga je veoma bitno da svaka aplikacija u svakom trenutku bude zaštićena od različitih tipova napada i pokušaja krađe osetljivih podataka, i na taj način zadobije i zadrži poverenje svojih korisnika.

Važan deo bezbednosti veb aplikacija jeste koncept polise zajedničkog porekla (eng. *same-origin policy*, *SOP*) [1]. Zahvaljujući ovom mehanizmu, skripte jedne veb stranice mogu da pristupe podacima druge veb stranice samo ako su istog porekla. Dva *URL*-a su istog porekla ako su im protokol, host i port identični. Na ovaj način sprečeno je da napadači preko svojih zlonamernih veb aplikacija dođu u posed osetljivih podataka smeštenih na nekom drugom veb sajtu. Zbog toga su napadači morali da osmisle nove načine kako mogu doći do korisničkih podataka, a da pritom zaobiđu polisu zajedničkog porekla.

XSS [2][3] je jedan od bezbednosnih propusta koji zaobilazi polisu zajedničkog porekla. Jedan je od retkih napada koji se iznova nalaze na *OWASP*-ovoj godišnjoj listi top 10 bezbednosnih propusta [4] i gotovo da ne postoji veliki veb sajt koji u nekom trenutku nije bio ranjiv na ovaj napad. *XSS* podrazumeva umetanje klijentskih skripti u ranjivu aplikaciju, koje su kasnije dostupne korisniku nakon učitavanja određenih veb stranica te aplikacije. *XSS* bezbednosni propust je i dalje popularan i zastupljen na velikom broju veb aplikacija. Razlog tome često može biti neiskustvo, neupućenost i neobazrivost programera koji izrađuju veb aplikacije, kao i nedostatak testiranja aplikacija na propuste prilikom svake velike izmene ili nadogradnje aplikacije.

Katarina Simić je student master studija na Elektrotehničkom fakultetu, Univerziteta u Beogradu, Bul. kralja Aleksandra 73, 11120 Beograd, Srbija (e-mail: sk193473m@etf.bg.ac.rs).

Žarko Stanisavljević radi na Elektrotehničkom fakultetu, Univerziteta u Beogradu, Bul. kralja Aleksandra 73, 11120 Beograd, Srbija (telefon: +381-11-3218-484; e-mail: zarko.stanisavljevic@etf.bg.ac.rs).

U ovom radu prikazana je implementacija i način korišćenja ranjive veb aplikacije, na kojoj je moguće na određenim mestima umetnuti zlonamerne skripte i izvršiti neku od zlonamernih akcija na štetu regularnog korisnika. Cilj ove aplikacije je jednostavna demonstracija nekih od najčešće primenjenih i praktičnih *XSS* napada, koja bi na taj način pomogla korisniku da bolje razume kada i kako ti napadi mogu da se dese, kao i na kojim mestima u aplikaciji. Nakon korišćenja aplikacije, korisnik bi trebao da razume osnovne koncepte *XSS* napada, kao i da bude u stanju da primeni odgovarajuće mere zaštite, koristeći naučeno, prilikom izrade sopstvene aplikacije.

U drugom poglavlju se opisuju detalji *XSS* sigurnosnog propusta. U trećem poglavlju je prikazan razvoj aplikacije koji se koristi u demonstrativne svrhe *XSS* napada, uz detaljan opis korišćenih tehnologija. U četvrtom poglavlju je opisan rad aplikacije i način korišćenja aplikacije. U petom poglavlju je dat zaključak.

II. *XSS*

Pojavom *JavaScript* programskog jezika sredinom devedesetih godina prošlog veka omogućen je veliki napredak u izradi veb aplikacija, koje su sada mogle biti i interaktivne. Ali, pored svih dobrih i interesantnih mogućnosti koje su sada bile dostupne, pojavile su se i one loše koje mogu uticati negativno po korisnika, poput *XSS* napada. Prvobitno se *XSS* napadu nije pridavalo mnogo pažnje, jer su serveri bili izazovnija i interesantnija meta napadačima. Ali tokom godina situacija se preokrenula. Serveri su vremenom postajali mnogo zaštićeniji nego ranije i bilo je sve teže probiti njihovu zaštitu. Uvidelo se i da serveri nisu bili neophodni za izvršavanje napada sa klijentske strane. Pojavljivali su se različiti pretraživači koji izvršavaju klijentski kod, svaki sa svojim propustima u zavisnosti od verzije, što je programerima dodatno otežavalo posao zaštite. Sa druge strane, programeri zbog manjka vremena ili budžeta, kao i manjka iskustva i znanja, ne posvećuju dovoljno pažnje bezbednosnim propustima, te ih je veoma lako i napraviti. Zbog svega ovoga se *XSS* danas smatra za jedan od najopasnijih i najučestalijih napada. Dve trećine svih veb aplikacija imaju *XSS* propuste u sebi, i svaka velika i popularna aplikacija je u nekom trenutku imala ovaj propust.

Primarni cilj napadača jesu korisnici ranjivih aplikacija. Napadi najčešće podrazumevaju krađu sesije, preuzimanje osetljivih podataka, izvršavanje nedozvoljenih akcija u ime korisnika, dostavljanje zlonamernih softvera korisniku (eng. *malware*), pa čak i narušavanje zaštite aplikacije od drugih napada. Osim što ovi napadi oštećuju same korisnike, mogu da imaju i destruktivne posledice po samu aplikaciju, ponajviše zbog gubitka poverenja od strane korisnika ako aplikacija zahteva visok nivo bezbednosti zbog veoma osetljivih i bitnih korisnikovih podataka podeljenih sa tom aplikacijom.

Postoji nekoliko varijacija *XSS* napada, koje se mogu podeliti u tri glavna tipa: reflektujući, snimljeni i *DOM* bazirani *XSS* napad.

A. Reflektujući XSS napad

Reflektujući *XSS* napad (eng. *Reflected XSS attack*) [5] je najzastupljeniji od sva tri tipa. Pod ovim napadom se podrazumeva da se umetnuta, zlonamerna skripta pošalje na server kao deo zahteva i zatim odmah reflektuje u korisnikovom pretraživaču u vidu odgovora koji sadrži tu skriptu. Dakle, podatak poslat serveru je vraćen i prikazan na stranici bez ikakve prethodne provere tog podatka. Da bi uspešno sproveo ovaj napad, napadač prvo mora da osmisli *URL* koji će sadržati zlonamernu skriptu. Zatim će taj *URL* proslediti korisnicima na lukav način. Ako neko od korisnika ništa ne posumnja, zahtevaće *URL* od aplikacije i server će vratiti odgovor koji će sadržati i napadačevu skriptu. Korisnikov pretraživač će sada, između ostalog, izvršiti i napadačevu skriptu i ukradeni podaci se šalju na napadačev server i postaju mu dostupni. Ovakva vrsta napada se izvršava samo ako korisnik otvori napadačev *URL*, te je tako ovaj napad često jednokratan. Da bi napadač naveo korisnika da nasedne i otvori sastavljen *URL*, mora da se posluži lukavim trikovima. Ako su mu meta individualne osobe, napadač će im zamaskirani *URL* upotpunjen uverljivom porukom poslati direkno (npr. preko *e-mail* poruke), tako da korisnik poželi da taj *URL* i otvori. Ako mu je meta veći broj ljudi, zamaskirani *URL* će postaviti na nekim drugim veb stranicama u vidu linka, te će čekati da neko taj link i otvori.

B. Snimljeni XSS napad

Snimljeni XSS napad (eng. *Stored XSS attack*) [6] je najopasniji tip *XSS* napada, zato što može da ima značajnije posledice po veći broj korisnika. Za razliku od reflektujućeg napada, zlonamerna skripta se čuva na serveru, tako da će se svakom korisniku, koji od servera bude zahtevao stranicu sa umetnutom skriptom, u pretraživaču ta skripta i izvršiti. Ovi napadi su najčešći na sajtovima gde korisnici imaju vid međusobne komunikacije (forumi, komentari, pitanja korisnika, itd.). Za izvršavanje ovog napada napadač ne mora da podmeće korisnicima direktno sastavljen *URL*, već je dovoljno da sam umetne skriptu u aplikaciji. Pošto će ona biti sačuvana na serveru, biće dostupna svakom korisniku te stranice. Ovaj napad će najverovatnije dati bolje rezultate u odnosu na reflektujući, jer da bi napadač ukrao korisnikove osetljive podatke, korisnik u najvećem broju slučajeva mora biti ulogovan, što će verovatno i biti slučaj prilikom izvršavanja snimljenog *XSS* napada, a manje verovatan slučaj prilikom reflektujućeg. Snimljenim napadom je veća i verovatnoća da napadačeva žrtva bude administrator napadnute veb aplikacije, što znači da cela aplikacija može biti komprimitovana i ugrožena.

C. DOM bazirani XSS napad

Za razliku od reflektujućeg i snimljenog *XSS* napada, za čije je izvršavanje neophodno vraćanje zlonamerne skripte sa servera, *DOM* (eng. *Document Object Model*) bazirani *XSS* napad [7] će se izvršiti bez da server vrati skriptu prosleđenu u *URL*-u. Ovo je moguće zato što *JavaScript* može pristupiti *DOM*-u i samim tim može dohvatiti i parametre *URL*-a. To znači da će zlonamerni kod biti preuzet sa *URL*-a i obrađen u *JavaScript* kodu. Ovakva vrsta napada ima više sličnosti sa reflektujućim *XSS* napadom nego snimljenim, jer zahteva od napadača da sastavljen *URL* na razne načine podmetne korisnicima, ali je zbog svoje prirode znatno opasniji. Razlika u odnosu na reflektujući *XSS* napad jeste što će *JavaScript* kod procesirati napadačev *URL*, pa samim tim i napadačevu skriptu umetnutu u *URL*, tako da bilo šta što će server vratiti kao odgovor nije važno prilikom ovog napada. Najveći problem kod ovog propusta jeste naći uzrok zbog kojeg nastaje, a pošto se taj uzrok može naći bilo gde u klijentskom kodu, programer bi morao dobro da poznaje projekat prilikom istrage.

D. Zaštita od XSS napada

Nakon upoznavanja sa XSS propustima i uviđanja njihovih mogućnosti i posledica po korisnike, naredni korak za programere bi bio da svoje aplikacije od istih i zaštite. S obzirom da je malo potrebno da se propusti naprave, neophodno je redovno testirati aplikacije na njih prilikom svake nadogradnje koda. Radi efikasnije zaštite aplikacija preporučljivo je primeniti metode poput validacije *input*-a [8], validacije *output*-a [9], konfiguracije aplikacije tako da vraća *Content-Security-Policy* zaglavlje [10], zabrane unosa korisničkih podataka na potencijalno opasnim mestima u okviru aplikacije [11], kao i korišćenja odgovarajućih zaglavlja odgovora koji bi mogli da detektuju *HTML* ili *JavaScript* kod u *HTTP* odgovorima, i samim tim spreče njihovo ubrizgavanje, poput *X-XSS-Protection* zaglavlja [12].

III. IMPLEMENTACIJA APLIKACIJE

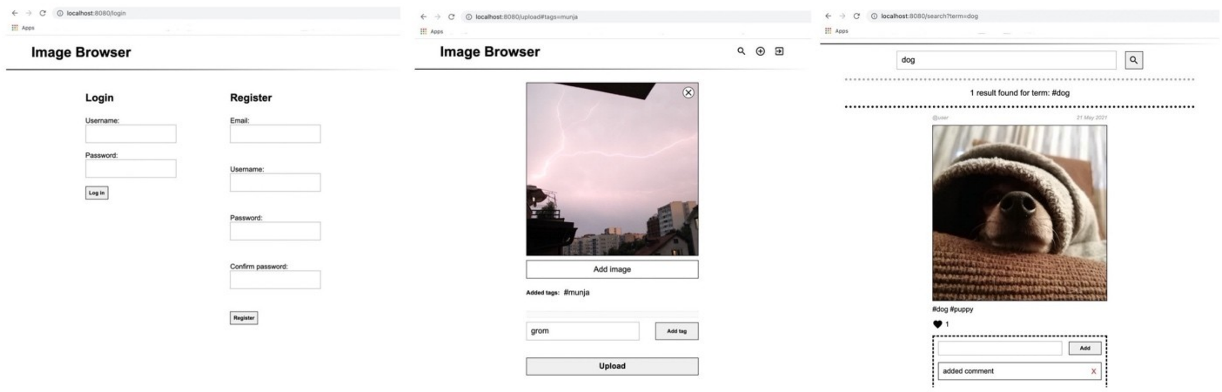
Implementirana aplikacija se sastoji iz dva dela. Prvi deo čini jednostavna, ali ranjiva veb aplikacija, koja služi za isprobavanje različitih *XSS* napada na različitim mestima u okviru te aplikacije. Drugi deo čini napadačev server, na koji pristižu ukradeni podaci nakon uspešno izvršenih napada.

A. Implementacija ranjive aplikacije

Prvi deo alata za učenje predstavlja jednostavnu veb aplikaciju za pretraživanje i dodavanje slika pod nazivom *ImageBrowser* (Sl. 1).

Na početku se od korisnika traži da se registruje ili uloguje na aplikaciju. Nakon što se korisnik uloguje, prikazuje mu se stranica sa porukom dobrodošlice. Nakon toga korisniku su dostupne dve različite stranice sa akcijama. Na prvoj stranici korisnik može da dodaje svoje slike uz koje ostavlja i određene tagove – reči koje služe za opisivanje slike. Na drugoj stranici korisnik može da, uz pomoć *input* polja, pretražuje slike na osnovu postojećih tagova, da pretražuje druge korisnike aplikacije dodajući simbol @ ispred korisničkog imena, kao i da prikazane slike komentariše i na njih reaguje.

Sama aplikacija ima minimalan skup potrebnih funkcionalnosti, ali i namerno napravljene *XSS* bezbednosne propuste na više mesta radi demonstracije nekih od *XSS* napada.

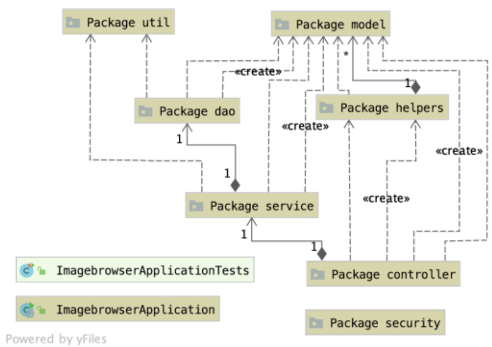


a) b) c)

Sl. 1. Izgled ranjive aplikacije u pretraživaču: a) početna stranica, b) stranica za dodavanje slika i c) stranica za pretraživanje slika

Da bi korisnik mogao da dodaje i pretražuje slike, kao i da ostavlja komentare i reakcije, određeni podaci moraju da se čuvaju u bazi podataka. Aplikacija koristi *H2 Java in-memory* bazu [13], koja omogućava da se prilikom svakog pokretanja aplikacije koristi inicijalno stanje podataka baze i da se sve dotadašnje izmene gube, što omogućava lakše testiranje aplikacije.

Za implementaciju servera je korišćen razvojni okvir *Spring* [14], kao i *SpringBoot* [15] koji koristi *Spring* kao podlogu. U projektu se koristi i *Apache Maven* [16], alat za izvršavanje `compile` i `build` naredbi *Java* koda. *Spring Initializer* [17] je projekat otvorenog koda (eng. *open source*), koji omogućava generisanje konfigurisanog *Spring* projekta uz odabir potrebnih zavisnosti (eng. *dependencies*). *Java* kod je grupisan po paketima, od kojih je okružujući *application.imagebrowser* u okviru koga se pored drugih nalazi i *ImagebrowserApplication,* glavna klasa aplikacije (Sl. 2).



Sl. 2. *UML* dijagram paketa unutar *application.imagebrowser* okružujućeg paketa

Unutar ovog paketa se nalazi nekoliko drugih paketa, koji sadrže sav potreban kod za uspostavljanje komunikacije i ispravan rad servera sa bazom podataka, obradu pristiglih zahteva na serveru i vraćanje odgovarajućih stranica i podataka koji se učitavaju na stranici sa kojima će korisnik interagovati.

Klijentska strana sadrži kod koji se izvršava u pretraživaču, i sadrži sve što korisnik može da vidi i sa čime može da interaguje. Najbitnije i najosnovnije tehnologije koje su korišćene prilikom izrade klijentskog dela aplikacije su *HTML* (eng. *Hyper Text Markup Language*), *CSS* (eng. *Cascading Style Sheets*) i *JavaScript*. U okviru posmatrane aplikacije se koriste *JSP* (eng. *Java Server Pages*) stranice [18] koje podržavaju dinamički sadržaj, što omogućava umetanje *Java* koda unutar *HTML* koda uz pomoć specijalnih *JSP* tagova. *JSP* komponenta predstavlja *servlet* koji ispunjava ulogu korisničkog interfejsa za *Java* veb aplikaciju. Radi olakšanog i preglednijeg fajla, uz *JSP* se koristi i biblioteka *JSTL* (eng. *The JSP Standard Template Library*) [19], koja omogućava da se *Java* kod zameni tagovima koji će raditi identičan posao. Fragmenti *JSP* koda koji se mogu koristiti na više mesta su smešteni u zasebne fajlove, koji se uz pomoć tagova učitavaju na određenoj *JSP* stranici. *JSP* fajlovi u aplikaciji koji čine stranice su:

- *login.jsp,* gde korisnik može da se uloguje ili registruje,
- *index.jsp*, koji predstavlja glavnu stranicu prikazanu korisniku nakon što se uloguje,
- *upload.jsp*, gde korisnik može da dodaje slike i tagove,
- *search.jsp*, gde korisnik može da pretražuje slike po tagovima ili drugim korisnicima, kao i
- *searchResults.jsp* i *noResults.jsp*, dve stranice koje server vraća kao rezultat *AJAX* [20] poziva, prva sa rezultatima, i druga sa informacijom da rezultata nema.

B. *Implementacija napadačevog servera*

Drugi deo aplikacije čini napadačev server, na kojem će se obrađivati pristigli zahtevi, tačnije prethodno dohvaćeni osetljivi podaci korisnika. Da bi zahtevi uopšte mogli da pristignu na server, napadač mora da sastavi *URL* koji će u sebi sadržati zlonamerni kod za dohvatanje podataka i za redirekciju. Nakon dohvatanja podataka, napadač može da ponovo izvrši redirekciju nazad ka napadnutoj aplikaciji tako da korisnik ni ne posumnja da je bio napadnut.

U okviru ove aplikacije je napadačev server implementiran u *Node.js* [21] platformi koja predstavlja asinhrono *runtime* okruženje za *JavaScript* jezik, i koja omogućava stvaranje skalabilnih veb aplikacija, samim tim i izvršavanje *JavaScript* koda van pretraživača.

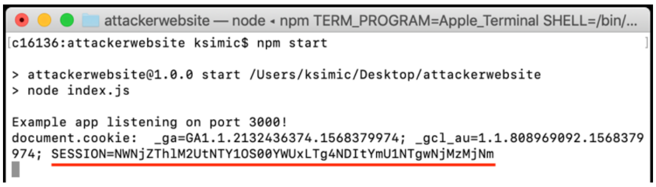
U slučaju napadačevog projekta je instaliran paket *Express* [22], koji predstavlja radni okvir za organizaciju aplikacije prema *MVC* arhitekturi. Pomoću *Express*-a se mogu na jednostavan način obrađivati pristigli zahtevi. Napravljen je jedan *JavaScript* fajl, *index.js*, u kojem se, prilikom izvršavanja koda, pokreće server koji osluškivanjem čeka na zahteve i obrađuje one koji su pristigli. Svaki zahtev se obrađuje tako da se u terminalu napadača gde je pokrenuta skripta ispišu pristigli podaci, a zatim po potrebi izvrši i redirekcija. Dokle god napadačev server radi, moći će da obrađuje pristigle zahteve.

IV. Način korišćenja aplikacije

U ovom poglavlju je dat prikaz nekoliko tipičnih scenarija *XSS* napada, gde je prvo objašnjen cilj napada, uz priložene zlonamerne skripte za ispunjenje tog cilja, a zatim je na kraju svakog primera dat savet za sprečavanje tog napada. Bitno je napomenuti da su ovakve vrste napada kažnjive zakonom svuda u svetu (npr. u Srbiji prema Krivičnom zakoniku Republike Srbije (Članovi 298 do 304a)) ukoliko se sprovode prema aplikacijama fizičkih i pravnih lica koja nisu upoznata i saglasna sa aktivnostima na proveri ranjivosti.

A. Krađa korisnikove sesije

Nakon što se korisnik uspešno uloguje na aplikaciju, server će poslati kolačić sesije preko *Set-Cookie* zaglavlja. Taj kolačić će se sada slati ka serveru uz svaki korisnikov zahtev. Zbog toga je kolačić sesije izuzetno osetljiv podatak, jer ako napadač nekako uspe da dođe do njegove vrednosti moći će da šalje zahteve ka serveru u ime oštećenog korisnika. *HttpOnly* predstavlja deo *Set-Cookie* zaglavlja u vidu *flag*-a. Ako je taj *flag* postavljen, to će sprečiti klijentske skripte da pristupe vrednostima kolačića. U slučaju da taj *flag* nije postavljen, krađu sesije je moguće izvesti. U slučaju ove aplikacije *flag* nije postavljen, tako da je moguće pristupiti vrednosti kolačića sesije u *JavaScript*-u preko *document.cookie* atributa (Sl. 3).



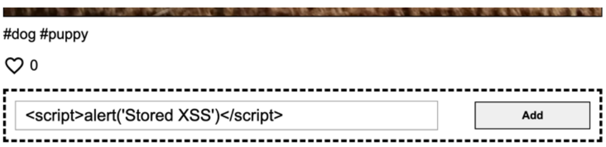
Sl. 3. Prikaz ukradenog kolačića sesije u terminalu napadača

1) Krađa kolačića sesije - reflektujući XSS napad

Na *search* stranici aplikacije postoji reflektujući *XSS* propust. Prilikom pretrage slika, uneti termin postaje *URL* parametar, i prilikom vraćanja rezultata od strane servera se vraća i pretražen termin koji se dodaje na stranicu. Validacija na tim osetljivim tačkama nije realizovana, samim tim korisnik umesto termina može da ukuca skriptu unutar `<script>` taga. Napadač takođe proverava na svojoj mašini da li *document.cookie* vraća njegovu tekuću sesiju. Nakon što utvrdi da vraća, napadač može da sastavi *URL* koji sadrži zlonamerni kod.

2) Krađa kolačića sesije - snimljeni XSS napad

U aplikaciji postoji i snimljeni *XSS* propust, tako da napadač može u komentare da ubacuje zlonamerni kod. Ako je situacija ista kao kod reflektujućeg *XSS* propusta, napadač sada može isti zlonamerni kod da doda kao komentar (Sl. 4). To znači da će svakom korisniku kojem se taj komentar bude učitao na stranici biti ukraden kolačić sesije. Ovo je suptilniji način za prevaru korisnika, te je veća verovatnoća da će korisnici biti prevareni ovom metodom nego da su kliknuli na *URL* primljen od napadača u slučaju reflektujućeg *XSS* napada.



Sl. 4. Trenutak dodavanja napadačeve skripte u komentar

3) Krađa kolačića sesije - DOM bazirani XSS napad

Na *upload* stranici aplikacije postoji *DOM* bazirani *XSS* propust. Korisnik može da dodaje tagove koji se pridodaju slici, i kako se neki tag doda on postaje deo *URL*-a u vidu `href` parametra. Problem je što se taj deo *URL*-a nikada ne šalje na server i ne obrađuje, ali se prilikom učitavanja stranice sa takvim *URL*-om tagovi automatski dodaju. To znači da se taj deo *URL*-a obradio negde u *JavaScript* kodu.

4) Način sprečavanja napada

Najlakši i najefikasniji način sprečavanja ovog napada jeste jednostavno podesiti *HttpOnly flag*, koji u tom slučaju sprečava klijentske skripte da dohvate podatke o kolačićima (Sl. 5). Na ovaj način *document.cookie* će uvek vratiti praznu vrednost i kolačić sesije će ostati bezbedan. Iako ovaj mehanizam odbrane od krađe kolačića funcioniše, to ne znači da napadač na istom mestu ne može da izvrši druge zlonamerne akcije. Ovaj *flag* se setuje na različite načine, u zavisnosti od korišćenog programskog jezika.

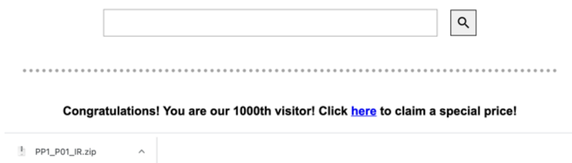
Sl. 5. Postavljanje *HTTPOnly flag*-a radi sprečavanja krađe kolačića sesije

B. Umetanje napadačevog koda na stranicu

U prethodnim primerima je data jednostavna skripta koja izvršava redirekciju i prosleđuje kolačić sesije napadaču, ali ako je primenjena navedena tehnika zaštite od tog napada, napadač mora da nađe drugi način da naškodi korisniku. Još jedna tehnika jeste umetanje *HTML* koda na stranicu, kojem se dodeljuju stilovi tako da izgleda kao da je zapravo deo stranice. Ako je kod dovoljno uverljiv, može da uveri korisnika da uradi određene akcije vođene tim kodom. U naredna dva primera se može videti kako umetanjem *HTML* koda napadač može da ukrade kredencijale korisnika, i kako može da navede korisnika da preuzme sumnjiv sadržaj na svoju mašinu.

1) Umetanje koda za preuzimanje sumnjivih fajlova

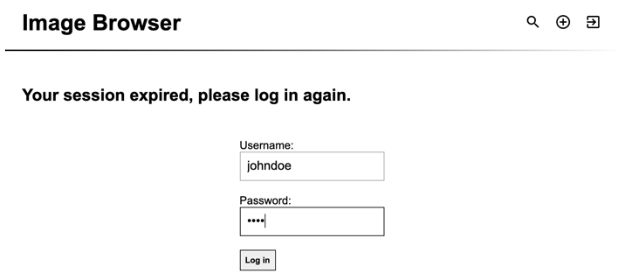
Ponekad napadaču nije samo cilj da ukrade korisnikove podatke, već i da ga navede da preuzme određeni fajl. To postiže umetanjem linka, na čiji klik se započinje preuzimanje nekog fajla. U zavisnosti od toga šta je cilj napadača, taj fajl može da ima nikakav ili razoran uticaj na korisnikovu mašinu. Dovoljno je samo sastaviti taj link dovoljno uverljivim da navede korisnika da klikne na njega, tako da će napadač i ovde uneti *inline* stilove, kao i uverljiv tekst (Sl. 6).



Sl. 6. Prikaz stranice za pretraživanje slike sa umetnutim linkom za preuzimanje sumnjivih fajlova

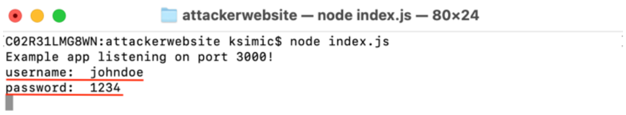
2) Phishing tehnika

Iako kolačići nakon primenjene zaštite ne mogu biti ukradeni, *XSS* propust i dalje postoji na istim mestima. Napadač uviđa da može da umetne skriptu koja sa stranice briše ceo *HTML* kod i zameni ga svojim. U ovom primeru je za brisanje i dodavanje koda korišćen *Jquery* (tačnije *.empty* i *.append* funkcije). Kod koji se dodaje predstavlja lažnu formu za unos korisničkog imena i šifre. Ta taktika umetanja ovakve vrste koda gde korisnik „dobrovoljno" ostavlja lične podatke napadaču se zove *phishing*. Napadač dodaje tekst da uveri korisnika da treba tu formu da popuni, na primer saopšti korisniku da mu je istekla sesija i da mora ponovo da se uloguje (Sl. 7).



Sl. 7. Prikaz stranice za pretraživanje slike sa umetnutom formom

Ako korisnik ništa ne posumnja, može uneti svoje lične podatke. Klikom na dugme za logovanje se zapravo izvršava redirekcija ka napadačevom serveru i štampaju se podaci u napadačevu konzolu i na kraju ponovo dešava redirekcija nazad ka aplikaciji. Napadač sad ima korisnikovo ime i šifru, što može da zloupotrebi na sličan način kao i prilikom krađe sesije (Sl. 8).



Sl. 8. Prikaz ukradenih kredencijala u terminalu napadača pomoću *phishing* tehnike

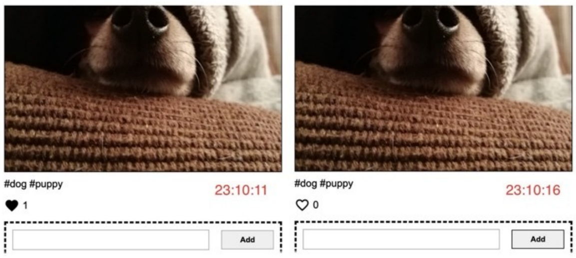
3) Način sprečavanja napada

U prethodnim primerima nije bilo potrebno raditi dodatnu validaciju zbog prirode napada. Ali u ovom slučaju bi bilo poželjno uraditi validaciju i *input*-a i *output*-a, na klijentskoj i na serverskoj strani. Prilikom pretrage se dešava *AJAX* poziv, pa je poželjno da se pre toga uradi sanitizacija *input*-a, najverovatnije uz pomoć regularnog izraza. U slučaju da *input* ne ispunjava zahteve, *AJAX* poziv se neće izvršiti i korisniku se ostavlja poruka da zna da je *input* polje bilo neispravno popunjeno. Sledeći korak je uraditi validaciju u kontroleru prilikom obrade zahteva i vraćanja odgovarajuće stranice. Tu bi najbolje rešenje bilo korišćenje gotovih metoda za zamenu (*escaping*) *HTML* koda. Ako je ipak potrebno dozvoliti određene tagove ili specijalne karaktere, metode moraju ručno da se pišu uz veliki oprez. Na primer, ako korisnik zameni `<script>` tag praznim stringom bez rekurzije, napadač može da sastavi skriptu sa *script* tagom `<scr<script>ipt>`. Takođe mora da se vodi računa da se obrade i *uppercase* i *lowercase* karakteri. Bitno je određene karaktere i stringove menjati u celoj skripti, a ne samo po prvom pojavljivanju. Zlonameran kod može biti prosleđen i kao atribut nekog dozvoljenog taga, pa je i za to potrebna validacija. Nakon što programer utvrdi sve šta mu je potrebno za validaciju na serveru, može da uradi i validaciju *output*-a. Najlakši način uraditi to u okviru ove aplikacije jeste koristeći *JSTL* tag `<c:out>`, jer ovaj tag omogućava *escaping* vraćenog koda. Na kraju se uz ove tehnike postiže visok nivo zaštite od ovakvog *XSS* napada.

C. Izvršavanje nedozvoljenih radnji u ime korisnika

Osim krađe podataka, napadač može pomoću zlonamernog koda i da izvrši neku radnju na stranici u ime korisnika. U tom slučaju nema potrebe da išta radi na svom serveru, već samo da pripremi kod koji će se izvršiti. Ako je cilj napadača da što više korisnika ošteti, to bi najbolje postigao uz snimljen *XSS* napad.

Skripta za ovaj primer je sastavljena tako da se u određenom vremenskom intervalu daju ili sklanjaju korisnikove reakcije na slike (Sl. 9). Iako ovaj napad nema veće posledice po korisnika, sam napad može da izazove nelagodnost i zbunjenost. Ovo je samo jedan primer izvršavanja nedozvoljenih akcija u skladu sa datim alatom, ali i u ovom slučaju zlonamerne skripte mogu da izazovu i znatno veće posledice, posebno ako korisnik nije ni svestan da se nešto desilo.



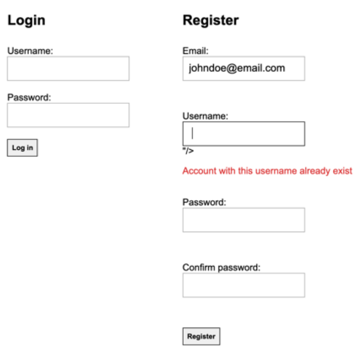
Sl. 9. Prikaz izvršavanja napadačeve skripte u ime korisnika u toku određenog vremenskog intervala

Iako je cilj napadača različit u odnosu na prethodni primer, metoda sprečavanja od *XSS* napada je ista. Potrebno je validirati *input* i *output*, sa klijentske i serverske strane. Poželjno je raditi validaciju na svim osetljivim mestima da bi se smanjio rizik od napada.

D. Keylogger

Još jedan način na koji napadač može da dođe do osetljivih korisnikovih podataka jeste da umetne skriptu koja napadačevom serveru prosleđuje karaktere koje korisnik unosi, karakter po karakter uz tačno vreme unosa. Ova tehnika nadgledanja pojedinačnih karaktera koje korisnik unosi se naziva *keylogger*, i najefikasnija je na stranicama na kojima se unose poverljivi podaci, poput broja kreditne kartice ili kredencijala. Skripta radi tako što pravi ograničen niz karaktera koji se šalje u određenom vremenskom intervalu samo ako je taj niz popunjen.

Na stranici za registraciju namerno je napravljen bezbednosni propust. Kada korisnik prilikom registracije unese već postojeće korisničko ime ili *e-mail*, prilikom osvežavanja stranice će ta informacija biti prisutna u *URL*-u i iz nje biti ispisana u odgovarajuće *input* polje (Sl. 10).



Sl. 10. Prikaz početne stranice u toku izvršavanja *keylogger XSS* napada

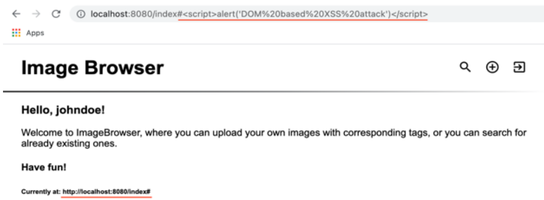
Napadač je otkrio da može svojim zlonamernim kodom da zatvori tag tog *input* polja, i dalje samo izvrši svoju skriptu. *Chrome*, *Safari* i *IE* pretraživači su se u ovom slučaju pokazali otpornim na napad zahvaljujući *X-XSS-Protection* zaglavlju odgovora. Svi korisnici koji se nalaze na ostalim pretraživačima će biti ranjivi. Napadač na ovaj način može da sazna kredencijale korisnika kroz individualne karaktere (Sl. 11).

Sl. 11. Prikaz ukradenih informacija u terminalu napadača tokom izvršavanja *keylogger XSS* napada

E. Jednostavan primer DOM baziranog XSS propusta

Preporučljivo je zaobići manipulaciju *DOM*-a u klijentskom kodu koliko god je to moguće da bi se smanjile šanse za *DOM* bazirani *XSS* napad. U ovom primeru se namerno na dnu svake stranice nalazi dekodovani *URL* tekuće stranice, koji se i dohvata i dekoduje u *JavaScript* kodu (Sl. 12). Ovde napadač takođe može da umetne skriptu koja će potpuno zaobići server i manipulisati kod koji se izvršava u *JavaScript*-u.

Dekodovani *URL* tekuće stranice se dohvata prilikom učitavanja te stranice u *JavaScript* kodu uz pomoć svojstva *document.URL*. Na ovaj način se vrši manipulacija *DOM* podataka, što otvara mogućnost da stranica bude ranjiva na *DOM* bazirani *XSS* napad. Eliminacijom tog koda problem bi u potpunosti nestao, ali ako je ipak potrebno taj kod i izvršiti, neophodno je uraditi validaciju. Tokom izvršavanja atributa *document.URL*, on se i dekoduje, te je preporučljivo ukloniti tu funkcionalnost. Ako informacija treba da bude dekodovana, treba izvršiti zamenu (eng. *escaping*) nepoželjnih karaktera koji se mogu naći u skripti sa odgovarajućom interpretacijom tog karaktera. Na ovaj način je sprečeno izvršavanje zlonamerne skripte na svim stranicama aplikacije.



Sl. 12. Prikaz ranjivosti stranice na kojoj se nalazi dekodovani *URL*

V. ZAKLJUČAK

U ovom radu predstavljena je jedna ranjiva aplikacija pomoću koje se može izučavati *XSS* sigurnosni propust. Aplikacijom su pokriveni različiti tipovi *XSS* napada. Korišćenjem aplikacije moguće je demonstrirati neke interesantne zloupotrebe ovih propusta. Najvažnije je da korisnici mogu na siguran način u zatvorenom okruženju detektovati propuste, a zatim koristeći tehnike zaštite ispraviti propuste i uveriti se da njihova rešenja ispravno rade. Aplikaciju je takođe moguće nadograditi po potrebi u budućnosti, radi dodavanja novih primera i propusta koji bi korisnicima dodatno omogućili testiranje i učenje o XSS napadu. Aplikacija je korišćena za izvođenje laboratorijskih vežbi na predmetu Zaštita računarskih sistema i mreža na Elektrotehničkom fakultetu u Beogradu, ali efekti korišćenja nisu izmereni usled uslova izvođenja nastave izazvanih pandemijom koronavirusa.

LITERATURA

[1] Same origin policy, pristupano 21.05.2021, https://developer.mozilla.org/en-US/docs/Web/Security/Same-origin_policy
[2] D. Stuttard, M. Pinto, “*The Web Application Hacker’s Handbook: Finding and Exploiting Security Flaws”,* second edition, John Wiley & Sons Inc, 2011
[3] Cross-site Scripting, pristupano 21.05.2021, https://owasp.org/www-community/attacks/xss/
[4] OWASP Top Ten Project, pristupano 21.05.2021, https://owasp.org/www-project-top-ten/
[5] Reflected XSS Attacks, pristupano 21.05.2021, https://owasp.org/www-community/attacks/xss/#reflected-xss-attacks
[6] Stored XSS Attacks, pristupano 21.05.2021, https://owasp.org/www-community/attacks/xss/#stored-xss-attacks
[7] DOM Based XSS, pristupano 21.05.2021, https://owasp.org/www-community/attacks/DOM_Based_XSS
[8] Input Validation Cheat Sheet, pristupano 21.05.2021, https://cheatsheetseries.owasp.org/cheatsheets/Input_Validation_Cheat_Sheet.html
[9] XSS Prevention Rules, pristupano 21.05.2021, https://cheatsheetseries.owasp.org/cheatsheets/Cross_Site_Scripting_Prevention_Cheat_Sheet.html#rule-1-html-encode-before-inserting-untrusted-data-into-html-element-content
[10] Content Security Policy (CSP), pristupano 21.05.2021, https://developer.mozilla.org/en-US/docs/Web/HTTP/CSP
[11] XSS Prevention Rules, pristupano 21.05.2021, https://cheatsheetseries.owasp.org/cheatsheets/Cross_Site_Scripting_Prevention_Cheat_Sheet.html#rule-0-never-insert-untrusted-data-except-in-allowed-locations
[12] X-XSS-Protection Header, pristupano 21.05.2021, https://cheatsheetseries.owasp.org/cheatsheets/Cross_Site_Scripting_Prevention_Cheat_Sheet.html#x-xss-protection-header
[13] H2 Database, pristupano 21.05.2021, https://www.h2database.com/html/main.html
[14] Spring, pristupano 21.05.2021, https://spring.io/
[15] Spring Boot, pristupano 21.05.2021, https://spring.io/projects/spring-boot
[16] Apache Maven, pristupano 21.05.2021, https://maven.apache.org/
[17] Spring Initializr, pristupano 21.05.2021, https://start.spring.io/
[18] JSP Tutorial, pristupano 21.05.2021, https://www.javatpoint.com/jsp-tutorial
[19] JSTL (JSP Standard Tag Library), pristupano 21.05.2021, https://www.javatpoint.com/jstl
[20] AJAX, pristupano 21.05.2021, https://developer.mozilla.org/en-US/docs/Web/Guide/AJAX
[21] Node.js, pristupano 21.05.2021, https://nodejs.org/en/
[22] Express, pristupano 21.05.2021, https://expressjs.com/

ABSTRACT

XSS (Cross-site scripting) is one of the most common vulnerabilities in web applications, despite the fact that there are many defense mechanisms against it that are available. This paper presents the implementation of a vulnerable application in which different types of XSS vulnerability can be demonstrated, along with the ways they can be misused, but also the ways they can be eliminated. The application can be used as an educational tool for software developer practical training in a closed and safe environment.

AN APPLICATION FOR DEMONSTRATION OF XSS VULNERABILITY

Katarina Simic, Zarko Stanisavljevic

SQLiTrainer - sistem za učenje o SQLi sigurnosnim propustima u aplikacijama

Đorđe Madić, Žarko Stanisavljević

***Apstrakt* — Sigurnosni propusti u aplikacijama koji nastaju prilikom njihovog razvoja i ostaju nedetektovani u produkcionom okruženju mogu dovesti do narušavanja integriteta, poverljivosti i dostupnosti takvih aplikacija. *SQLiTrainer* predstavlja skup ranjivih aplikacija kojima se mogu demonstrirati različite vrste *SQLi (eng. SQL injection)* ranjivosti. U radu je opisan način implementacije *SQLiTrainer* sistema i dati su primeri na koji način se sistem može iskoristiti za praktičnu obuku programera. Sistem je uspešno korišćen za izvođenje laboratorijskih vežbi na predmetu Zaštita računarskih sistema i mreža na Elektrotehničkom fakultetu u Beogradu.**

***Ključne reči* — SQLi, sigurnosni propusti, razvoj bezbednog softvera.**

I. Uvod

Razvoj bezbednog softvera podrazumeva postojanje svesti kod programera o potencijalnim problemima, a zatim i primenu čitavog seta dobrih praksi, kao i automatizovanih alata tokom procesa razvoja softvera. Aplikacije kod kojih postoje sigurnosne ranjivosti koje se mogu zloupotrebiti mogu dovesti do štete kako za korisnike takvih aplikacija, tako i za njihove autore.

Jedan od sigurnosnih propusta koji je često zastupljen u veb aplikacijama je *SQL injection* [1], kod koga se na različite načine na nepredviđen način mogu umetnuti naredbe koje mogu narušiti integritet, poverljivost i dostupnost baza podataka. Ovaj propust se već duži niz godina nalazi na top listama najčešćih sigurnosnih propusta u aplikacijama koje objavljuju organizacije kao što je *OWASP (Open Web Application Security Project)* [2].

U opštem slučaju nije jednostavno omogućiti programerima da kroz praktičan rad unaprede svoje znanje o ovakvim problemima, jer to podrazumeva izučavanje različitih mehanizama kojima se narušava informaciona bezbednost aplikacija na kojima se primenjuju. Primena ovih mehanizama kada se izvršavaju prema sistemima fizičkih i pravnih lica koja nisu upoznata i saglasna sa aktivnostima na proveri ranjivosti i testiranju upada u njihove sisteme je kažnjiva svuda u svetu (npr. u Srbiji prema Krivičnom zakoniku Republike Srbije (Članovi 298 do 304a)). Iz tog razloga postoje različiti sistemi koji omogućavaju svojim korisnicima da u zatvorenom okruženju na praktičan način obave obuku, a da ne prekrše zakon [3-5]. U ovom radu prikazan je jedan novi sistem za učenje o *SQLi* sigurnosnim propustima u aplikacijama.

Đorđe Madić radi u kompaniji Zuehlke Engineering, Bul. Milutina Milankovića 1i, 11070 Novi Beograd, Srbija (e-mail: djordje.madic@zuehlke.com).

U trenutku rada na ovom istraživanju Đorđe Madić je bio student master studija na Elektrotehničkom fakultetu, Univerziteta u Beogradu.

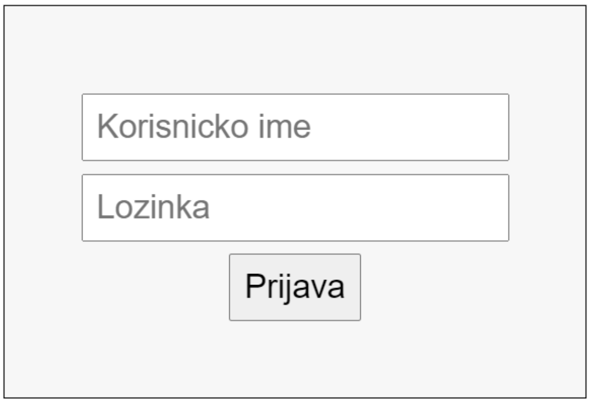
Žarko Stanisavljević radi na Elektrotehničkom fakultetu, Univerziteta u Beogradu, Bul. kralja Aleksandra 73, 11120 Beograd, Srbija (telefon: +381-11-3218-484; e-mail: zarko.stanisavljevic@etf.bg.ac.rs).

U drugom poglavlju su na primeru opisani uzrok i koraci kod izvođenja *SQLi* napada. U trećem poglavlju prikazan je način korišćenja realizovanog sistema na primeru jedne laboratorijske vežbe. U četvrtom poglavlju prikazana je implementacija laboratorijskih vežbi. U petom poglavlju dat je zaključak.

II. *SQL Injection*

OWASP definiše *injection* kao „slanje nepouzdanih podataka interpreteru kao deo komande ili zahteva“. *SQL injection* je tip napada koji se koristi za neovlašćeni pristup *SQL* bazi podataka koju aplikacija koristi. Napadač može čitati osetljive podatke, menjati njihovu strukturu, a u nekim slučajevima i izvršavati komande nad operativnim sistemom baze. Uzrok postojanja propusta je što aplikacija dozvoljava da korisnikov unos učestvuje u kreiranju *SQL* upita, omogućavajući mu da modifikuje originalni upit u svoju korist. U nastavku je na primeru *SQLi Login Bypass* napada objašnjen postupak izvršavanja napada.

SQLi Login Bypass počinje na stranici za prijavu korisnika, kao na Sl. 1. Cilj napada je, u bukvalnom prevodu sa engleskog, „zaobići prijavu“, odnosno prijaviti se ne znajući ni jedno korisničko ime ni lozinku.



Sl. 1. Stranica za prijavu korisnika

Prvi korak napada je analiza aplikacije. Ideja je pronaći ulazne podatke aplikacije koji se potencijalno koriste za kreiranje *SQL* upita. Ti podaci su kandidati da „nose“ napadački upit. U primeru prijave korisnika postoje dva ulazna podatka, korisničko ime i lozinka.

U narednom koraku potrebno je pretpostaviti kako izgleda *SQL* upit i odabrati ulazne podatke za napad. Na primer, upit za prijavu može biti sledeći:

```
SELECT * FROM korisnik
WHERE korisnicko_ime='$korisnicko_ime'
AND lozinka='$lozinka'
```

Delovi upita *$korisnicko_ime* i *$lozinka* su ulazni podaci sa korisničkog interfejsa, a napad se, na primer, može izvršiti kroz *$korisnicko_ime*.

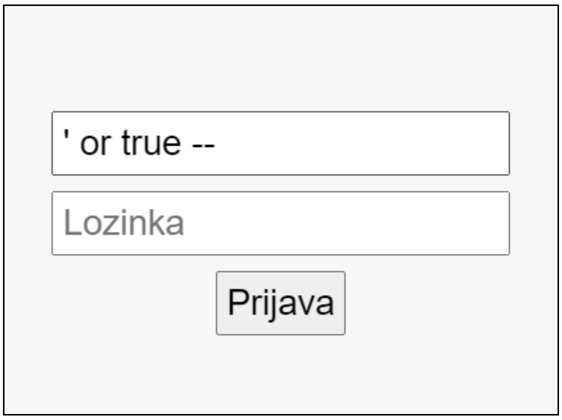
Napad se nastavlja na korisničkom interfejsu. U najčešćem slučaju tekstualni unos u polje za korisničko ime biće tretiran kao podatak, odnosno neće biti interpretiran kao komanda od strane baze podataka. Kako je korisničko ime tekstualnog tipa, u upitu se koristi apostrof za označavanje početka i kraja tekstualnog podatka. U slučaju da se u polje za korisničko ime unese apostrof, karakteri koji prethode tretiraće se kao podatak, dok će se oni koji slede tretirati kao komanda.

Koristeći ovu činjenicu napadač ima priliku da modifikuje originalni upit. Pod pretpostavkom da je prijava korisnika uspešna ukoliko upit vrati bar jedan rezultat, napad se može izvršiti kao na Sl. 2. što rezultuje sledećim upitom:

```
SELECT * FROM korisnik
WHERE korisnicko_ime='' or true --' and
password=''
```

Rezultat upita su svi redovi tabele „korisnik". Simbol „--" predstavlja oznaku za komentar čime se ignoriše deo upita nakon korisničkog imena.

Uslov „*or true*" čini da svaki red bude deo rezultata, ignorišući korisničko ime.



Sl. 2. *SQLi* napad na slučaju korišćenja prijave korisnika

Osnovu prevencije *SQLi* sigurnosnog propusta čine parametrizovani upiti (eng. *Parametrized Statement*, *Prepared Statement*). Oni omogućavaju da se bazi prvo prosledi šablon upita koji će se koristiti, a zatim za svako izvršavanje šablona i konkretni podaci. Korišćenje parametrizovanih upita garantuje da konkretni podaci neće biti interpretirani, čime se eliminiše *SQLi* sigurnosni propust. U nastavku je primer korišćenja parametrizovanog upita za pronalazak korisnika sa određenim korisničkim imenom u programskom jeziku *Java*:

```
String query = "SELECT * FROM korisnik WHERE
korisnicko_ime = ?";
PreparedStatement preparedStatement =
connection.prepareStatement(query);
preparedStatement.setString(1, "petar");
ResultSet results =
preparedStatement.executeQuery();
```

III. Primer korišćenja *SQLiTrainer* sistema

Primer korišćenja realizovanog sistema biće dat kroz prikaz jedne od laboratorijskih vežbi, dok će način upotrebe sistema u nastavi biti prikazan na primeru predmeta Zaštita računarskih sistema i mreža (ZRM) [6] na Elektrotehničkom fakultetu u Beogradu (ETF).

A. *Laboratorijska vežba Megatron*

Laboratorijska vežba *Megatron* zasniva se na aplikaciji koja predstavlja veb prodavnicu kompjuterske opreme. Vežba počinje na stranici za pretragu proizvoda. Pretraga se vrši po nazivu, gde je moguće uneti i samo deo naziva proizvoda. Rezultat pretrage prikazuje se u vidu tabele, sa kolonama za naziv i cenu proizvoda, kao na Sl. 3.

Sl. 3 Pretraga prozvoda

Cilj laboratorijske vežbe je pronaći korisničko ime i lozinku svih korisnika aplikacije. Boduju se i sledeće informacije:

- Naziv baze koju koristi aplikacija
- Verzija baze
- Nazivi tabela
- Nazivi kolona

Napad se izvršava kroz polje za pretragu proizvoda, dok se podaci koji su rezultat napada prikazuju u tabeli rezultata pretrage. Na primer, unosom sledećeg teksta u polje za pretragu dolazi se do naziva baze:

```
' and false union select 1, database() --
```

Sledećim unosom dolazi se do verzije baze:

```
' and false union select 1, h2version() --
```

Kako bi se došlo do korisničkog imena i lozinke svih korisnika aplikacije, prvo je potrebno pronaći naziv tabele koja sadrži korisnike. Sledećim unosom dolazi se do naziva svih tabela u bazi:

```
' and false union select 1, table_name from
information_schema.tables where
table_schema=database() --
```

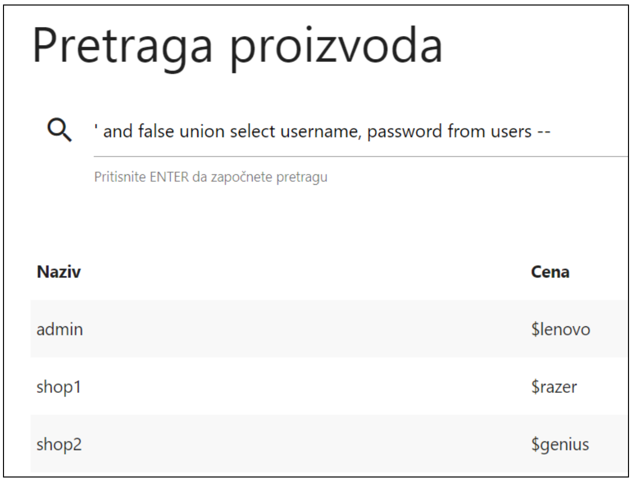
Iz prethodnog koraka saznaje se da je tabela sa korisnicima sačuvana pod imenom *users*. Sledećim unosom dolazi se do naziva svih kolona ove tabele:

```
' and false union select 1, column_name from
information_schema.columns where
table_name='users' --
```

Poslednji korak je definisanje upita koji prikazuje korisničko ime i lozinku svih korisnika:

```
' and false union select username, password
from users --
```

Tabela sa rezultatima sadržaće korisničko ime i lozinku svih korisnika, kao na Sl. 4.



Sl. 4 Korisničko ime i lozinka svih korisnika aplikacije

B. *Način upotrebe u nastavi*

ZRM je predmet master studija na Modulu za računarsku tehniku i informatiku ETF-a, koji je razvijen u okviru Erasmus+ KA2 projekta pod nazivom *Information Security Services Education in Serbia (ISSES)* [7]. Uzimajući u obzir probleme kod praktičnog izučavanja tema koje se obrađuju na predmetu, a koji su pomenuti u poglavlju I, za studente je napravljeno zatvoreno virtuelno laboratorijsko okruženje u okviru Laboratorije za informacionu bezbednost, koja je uspostavljena i opremljena u okviru istog (*ISSES*) projekta.

Svaki student ima sopstveno virtuelno laboratorijsko okruženje kome pristupa korišćenjem *VPN* veze. Za različite teme koje se obrađuju na predmetu koriste se različite konfiguracije virtuelnih laboratorijskih okruženja. U slučaju laboratorijskih vežbi u kojima se koristi *SQLiTrainer* sistem laboratorijsko okruženje se sastoji od jedne virtuelne mašine na kojoj je pokrenut *Ubuntu Linux* i na kojoj je pokrenut *SQLiTrainer* sistem. Studenti mogu da pristupe aplikacijama *SQLiTrainer* sistema iz svojih pretraživača korišćenjem *VPN* veze.

SQLiTrainer sistem se koristi za izvođenje laboratorijiskih vežbi, ali i kao deo finalnog praktičnog ispita na predmetu. Zahvaljujući načinu implementacije sistema, uz minimalne izmene u kodu, moguće je jednostavno izmeniti svaku od aplikacija tako da se dobiju drugačiji problemi sa istom tematikom, čime je omogućeno da se isti sistem iskoristi i prilikom obučavanja studenata, ali i prilikom provere njihovog znanja.

Još jedan važan aspekt, kada je u pitanju upotreba sistema, jeste i jednostavnost instalacije i konfiguracije. Prilikom konfigurisanja laboratorijskog okruženja za izvođenje laboratorijskih vežbi potrebno je pokrenuti više aplikacija istovremeno. Način implementacije *SQLiTrainer* sistema omogućava da se svaka aplikacija može pokrenuti na različitom portu, čime se prethodno postiže na jednostavan način. Kada je u pitanju instalacija, jedno rešenje je da se aplikacije iskopiraju na svaku virtuelnu mašinu i pokrenu odgovarajućim komandama. Ovo rešenje je vremenski zahtevno i nepraktično kada postoji veliki broj studenata na predmetu, a samim tim i veliki broj laboratorijskih okruženja koje je potrebno pripremiti. Način implementacije *SQLiTrainer* sistema dozvoljava da se iskoristi neki od alata za automatizaciju, kao što je na primer *Ansible* [8], čime se prethodni problem efikasno rešava na taj način što se napišu odgovarajuće skripte za ovaj alat kojima se prethodno manuelni posao kopiranja i pokretanja aplikacija u potpunosti automatizuje.

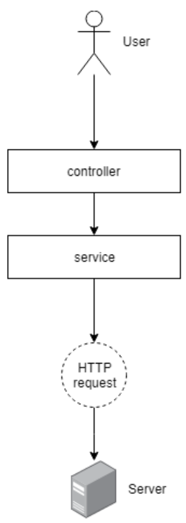
Opisani sistem je korišćen u nastavi u dve uzastopne školske godine 2019/2020 i 2020/2021. Prve školske godine finalni praktični ispit uspešno je savladalo 67% studenata (22/33), dok je u drugoj školskoj godini uspešno bilo 65% studenata (53/81).

IV. Implementacija *SQLiTrainer* sistema

Sistem čine četiri aplikacije identične strukture. Dve demonstriraju *Union-based SQLi*, pravolinijski i jednostavan napad gde se u kratkim iteracijama otkriva sve više podataka iz baze. Sledeća demonstrira *Blind SQLi* za koju je specifično da se podaci iz baze nikada ne prikazuju napadaču, i spada u teže napade za manuelno izvršavanje. Poslednja demonstrira *SQLi Login Bypass*, gde je cilj napadača da uspešno izvrši korisničku prijavu bez prethodnog poznavanja bilo kog korisničkog imena ili lozinke. U ovoj aplikaciji student se upoznaje sa upotrebom *HTTP Proxy* server, kao alata u izvršavanju *SQLi* napada.

Sistem je implementiran kao skup *Java* veb aplikacija koristeći *Spring Boot* [9] i *h2* [10] *in-memory* bazu podataka. Korisnički interfejs aplikacija kreiran je koristeći *HTML*, *CSS* i *JavaScript*. Za kreiranje lepšeg korisničkog interfejsa korišćen je *MaterializeCSS* [11], dok *AngularJS* [12] pojednostavljuje pisanje koda za interakciju sa korisnikom i *HTTP* (eng. *Hypertext Transfer Protocol*) komunikaciju sa serverom.

Korisnički interfejs i serverska aplikacija mogu se posmatrati kao dve odvojene aplikacije koje komuniciraju preko *HTTP*-a.



Sl. 5 Životni ciklus zahteva u aplikaciji korisničkog interfejsa

Korisnički interfejs realizovan je kao *Single Page Application* (*SPA*). Srž aplikacije čine *HTML* stranica *index.html* i *JavaScript* kod *app.js*. Resursi aplikacije organizovani su u posebne direktorijume:

- *css* – sadrži *CSS* biblioteke i definicije stilova specifičnih za aplikaciju,
- *images* – sadrži fotografije korišćene na korisničkom interfejsu i
- *js* – sadrži *JavaScript* biblioteke i kod aplikacije.

Organizacija koda je slojevita i definiše dva sloja:

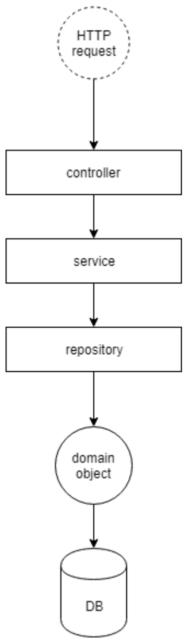
- *controller* – kod koji obrađuje korisničke akcije i
- *service* – kod za komunikaciju sa serverskom aplikacijom.

Sloj *controller* je viši sloj i zavisan od sloja *service*, a korisnički zahtevi prolaze kroz oba sloja aplikacije. Na Sl. 5 prikazano je kako korisnički zahtev putuje kroz aplikaciju korisničkog interfejsa i do serverske aplikacije.

Organizcija serverskog koda je takođe slojevita, gde su slojevi aplikacije predstavljeni sledećim *Java* paketima:

- *controller* – klase za razmenu podataka sa aplikacijom korisničkog interfejsa,
- *service* – klase koje izvršavaju poslovnu logiku aplikacije i
- *repository* – klase za čitanje i upis domenskih objekata u bazu podataka.

Pored navedenih paketa postoji i paket *domain* koji sadrži klase koje predstavljaju domenske entitete.



Sl. 6 Životni ciklus zahteva u serverskoj aplikaciji

Zahtevi koji dolaze sa korisničkog interfejsa prolaze kroz sve slojeve aplikacije. Način razmene podataka između slojeva prikazan je na Sl. 6 na primeru zahteva koji na kraju rezultuje upisom u bazu podataka. Svaki sloj aplikacije zavisan je od sledećeg (nižeg) sloja, a često su svi paketi aplikacije zavisni od domenskog. U nekim implementacijama izostavljen je servisni sloj, jer ne bi sadržao nikakvu logiku, već bi samo prosleđivao podatke sledećem sloju.

Svaka instanca aplikacije poseduje svoju instancu baze, koja se pokreće zajedno sa aplikacijom i čuva podatke u memoriji. Kod serverske aplikacije prate *SQL* skripte koje se izvšavaju nad bazom prilikom pokretanja aplikacije, kako bi pri svakom pokretanju aplikacije stanje baze bilo identično. Skripte se nalaze u sledećim fajlovima:

- <u>*schema.sql*</u> – izvršava se prva i njena uloga je da kreira relacionu šemu baze i
- <u>*data.sql*</u> – popunjava bazu podacima.

Prilikom zaustavljanja aplikacije zaustavlja se i baza podataka, a podaci iz nje trajno nestaju.

Varijacije problema laboratorijskih vežbi mogu se kreirati na više načina. Najjednostavniji je izmeniti *SQL* skripte čime se menja početno stanje baze podataka. Kompleksnije izmene, poput izmene imena kolona i tabela, zahtevaju i manje izmene u kodu aplikacije. Navedene skripte se mogu pokrenuti i nad nekom drugom bazom podaka, na primer *PostgreSQL* ili *MySQL*, za šta je dovoljno u konfiguracionom fajlu aplikacije navesti parametre za konekciju. Korišćenje različitih baza podataka povećava kompleksnost zadatka jer koriste različite dijalekte *SQL* jezika.

V. Zaključak

U ovom radu predstavljen je jedan novi sistem za učenje o *SQLi* sigurnosnim propustima u aplikacijama. *SQLiTrainer* služi za praktičnu obuku programera u oblasti razvoja bezbednog softvera. Realizovan je kao skup ranjivih aplikacija kojima se mogu demonstrirati različiti tipovi *SQLi* propusta koji se mogu javiti u aplikacijama. Omogućava proveru postojanja propusta u sigurnom okruženju, kao i učenje tehnika kojima se mogu otkloniti uočeni propusti. U radu je prikazan opis *SQLiTrainer* sistema i način njegovog korišćenja. U budućnosti se planira dodavanje skupa vežbi u kojima će studenti biti u prilici da isprave propuste koji postoje u aplikacijama.

Zahvalnica

Autori žele da se zahvale master inž. Adrianu Milakoviću i prof. dr Pavlu Vuletiću na pomoći prilikom uvođenja sistema na laboratorijske vežbe na predmetu Zaštita računarskih sistema i mreža.

Literatura

[1] „*SQL injection*,“ Dostupno na: https://owasp.org/www-community/attacks/SQL_Injection, poslednji put pristupano: maj 2021.

[2] „*OWASP*,“ Dostupno na: https://owasp.org/, poslednji put pristupano: maj 2021.

[3] „*Avatao*,“ Dostupno na: https://avatao.com/, poslednji put pristupano: maj 2021.

[4] „*Juice Shop*,“ Dostupno na: https://owasp.org/www-project-juice-shop/, poslednji put pristupano: maj 2021.

[5] „*Security Idiots*,“ Dostupno na: http://www.securityidiots.com/, poslednji put pristupano: maj 2021.

[6] „Zaštita računarskih sistema i mreža,“ Dostupno na: https://www.etf.bg.ac.rs/fis/karton_predmeta/13M111ZRM-2019, poslednji put pristupano: maj 2021.

[7] „*Information Security Services Education in Serbia*,“ Dostupno na: https://isses.etf.bg.ac.rs/, poslednji put pristupano: maj 2021.

[8] „*Ansible*,“ Dostupno na: https://www.ansible.com/, poslednji put pristupano: maj 2021.

[9] „*Spring Boot*,“ Dostupno na: https://spring.io/projects/spring-boot, poslednji put pristupano: maj 2021.

[10] „*H2 database*,“ Dostupno na: https://www.h2database.com /html/main.html, poslednji put pristupano: maj 2021.

[11] „*Materialize CSS*,“ Dostupno na: https://materializecss.com/, poslednji put pristupano: maj 2021.

[12] „*AngularJS*,“ Dostupno na: https://angularjs.org/, poslednji put pristupano: maj 2021.

Abstract

During the application development process security vulnerabilities can occur and remain in application in production environment. These vulnerabilities can cause confidentiality, integrity and availability breaches. SQLiTrainer represents a set of vulnerable applications that can be used to demonstrate different types of SQLi vulnerabilities. Implementation of the SQLiTrainer system is given in the paper and the examples on how to use the system for programmer practical training is proposed. The system was successfully used for laboratory exercises at the Advanced System and Network Security course at the University of Belgrade, School of Electrical Engineering.

SQLITRAINER - SYSTEM FOR LEARNING ABOUT SQLI VULNERABILITY IN APPLICATIONS

Djordje Madic, Zarko Stanisavljevic

Jedno rješenje analize i prikaza kontrolnih tačaka definisanih podešavanjem AUTOSAR nadzornog časovnika

Ivana Tešević, Branko Milošević, Dejan Bokan i Bogdan Pavković

Apstrakt— **Razvojem automobilske industrije i softvera unutar nje kao tehnička posljedica javila se potreba za obaveznom integracijom zaštitnih mehanizama u ugrađenim jezgrima operativnog sistema. Jedan od osnovnih mehanizama za zaštitu sistema jeste nadzorni časovnik (eng. watchdog, WDG). Ova komponenta ima za cilj da nadgleda sve ostale komponente pokrenute od strane raspoređivača i time omogući bezbjedan rad sistema. Kako je probleme koje nadzorni časovnik prijavljuje relativno teško ispratiti i analizirati u stvarnom sistemu, došlo se do ideje da se oponaša rad komponente nadzornog časovnika na računaru sa istim ulaznim parametrima kao u živom sistemu. U ovom radu je dato rješenje za simulaciju mehanizama nadgledanja sistema definisane AUTOSAR arhitekture. Simulacijom je omogućeno da se minimalizuju odstupanja, predvide greške u sistemu i olakša sama analiza. Rad može doprinijeti bržem razvoju sistema jer omogućava da se prije implementacije predvide greške koje će se desiti u sistemu.**

Ključne riječi— AUTOSAR, WDG, WdgM, WdgIf, nadgledanje krajnjih rokova, logički nadzor, nadgledanje u realnom vremenu, najduže vrijeme izvršenja(WCET).

I. Uvod

AUTOMOBILSKA industrija je grana industrije koja se sveobuhvatno razvija u posljednjoj deceniji. Dizajn vozila u automobilskoj industriji tradicionalno se oslanja na diskretne hardverske komponente (elektronske upravljačke jedinice - ECU), sa vrlo malo potrebnog softvera. Sa poboljšanjem automobilske industrije, softver za nju se razvijao. Danas je softverski dio prevladao hardver[1][2] . Softverske komponente postale su komplikovanije i zahtjevnije od hardverskih komponenti. Danas automobili nude mnogo više mogućnosti, uključujući i autonomne funkcije pri vožnji[1]. Postoji pet nivoa automatizacije vožnje, dok je industrija trenutno na trećem nivou, očekujući da će dostići nivo četiri i pet do 2025. godine[3]. Vozači će moći bezbjedno da skrenu pažnju sa vožnje, npr. gledati film ili čitati knjigu.

Činjenica je da je sve više dobavljača u ovoj grani industrije, pa se pojavila potreba za standardizacijom proizvodnje softvera. Da bi se udovoljilo ovom zahtjevu, stvorena je platforma AUTOSAR (eng. *Automotive Open System Architecture*) [1].

Kako se ova industrija sve više širi i kako rastu softverski zahtjevi povećava se i potreba za raznim alatima za održavanje bezbjednosti sistema. Ovakvi sistemi moraju podlijegati raznim testovima i konstantno se nadgledati kako bi se u potpunosti otklonila mogućnost greške, jer i najmanja greška može imati fatalne posljedice.

Nadzorni časovnik je jedna od komponenti koja za cilj ima nadzor cijelog sistema. Ova komponenta kao takva sama po sebi mora imati maksimalni kvalitet koda i podlijegati najvećim provjerama. Bilo kakva greška primijećena od strane WDG komponente biće ispraćena reakcijom gašenja cijelog sistema. Ovakav vid zaštite u industriji otežava testiranje, predviđanje ali i pronalaženje greške u toku rada. Zato se javila potreba da se prikaže jedno rješenje za oponašanje sistema kako bi se moglo predvidjeti i upoznati sa greškama i načinima na koji dolazi do njih.

Ovaj rad prikazuje jedno rješenje analize i prikaza kontrolnih tačaka definisanih podešavanjem *AUTOSAR* nadzornog časovnika. Prikazaće se simulacija poremećaja u sistemu koji će nadzorni časovnik prepoznati pomoću nadzornih mehanizama. Namjerno izazivanje poremećaja i reakcija nadzornog časovnika na te poremećaje doprinijeće lakšem testiranju i predviđanju u stvarnom sistemu. Pomoću ovog rješenja nudi se mogućnost korišćenja stvarnih ulaznih parametara i testiranje raznih poremećaja i lanca događaja nakon namjerno izazvane greške. Mogućnost predviđanja i vizualni prikaz sistema nakon poremećaja jesu glavni doprinos ovog rada. Postojeća literatura na temu nadzornog časovnika[5][6] skoncentrisana je na unaprjeđenju mehanizama zaštite ili na načinu testiranja sistema i ispravnosti sprege nadzornog časovnika.

Drugo poglavlje će dati teorijske osnove o načinu rada svih modula, samom *AUTOSAR* standardu i vertikali nadzornog časovnika sa definisanim modulima.

U trećem poglavlju biće opisan način rada, pristup rješenju i podloga za nastavak i samu implementaciju.

U četvrtom poglavlju su opisani moduli, dato je programsko rješenje i sami postupci implementacije.

U petom poglavlju su prikazani rezultati rada, način testiranja, kao i svrha samog rješenja.

II. AUTOSAR standard

Osnovan 2003. godine, AUTOSAR predstavlja međunarodno razvojno partnerstvo stranaka iz automobilske industrije. Cilj ove saradnje bio je stvaranje i uspostavljanje otvorene i standardizovane softverske arhitekture za osnovne elektronske jedinice autonomnog vozila nazvane ECU.

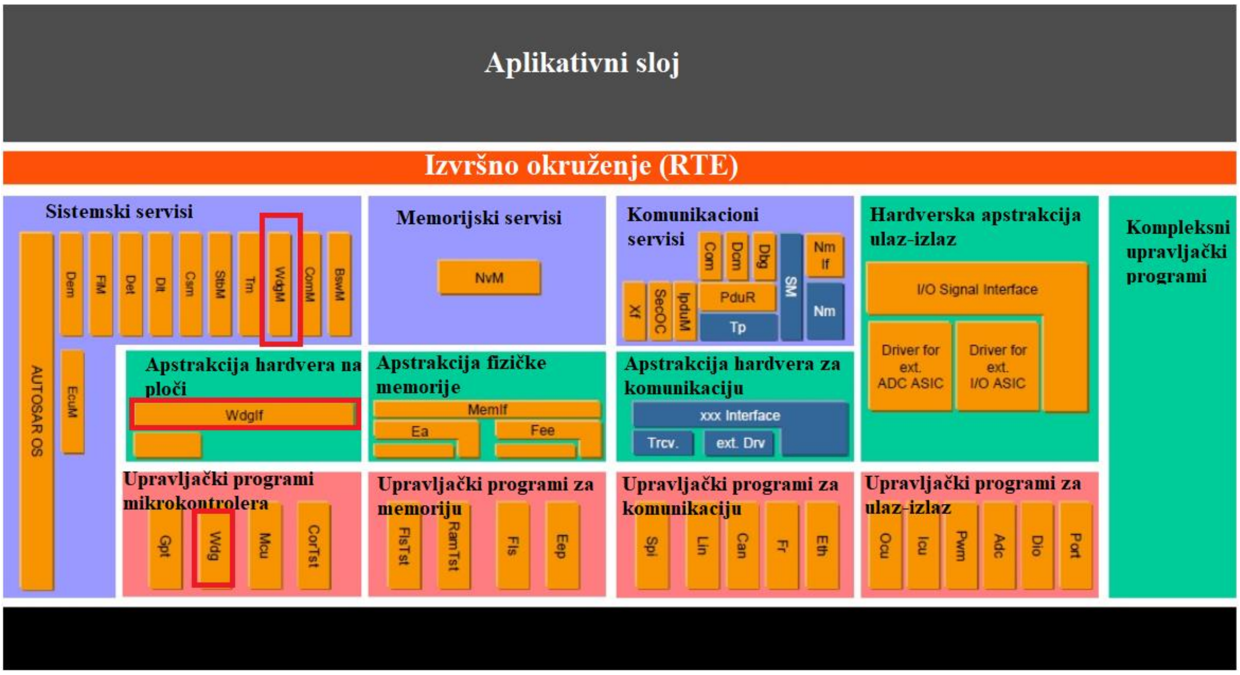
Ivana Tešević, RT-RK Institute for Computer Based Systems, Novi Sad, Srbija (e-mail: ivana.tesevic@rt-rk.com)

Branko Milošević, RT-RK Institute for Computer Based Systems, Novi Sad, Srbija (e-mail: branko.milosevic@rt-rk.com).

Bogdan Pavković, RT-RK Institute for Computer Based Systems, Novi Sad, Srbija (e-mail: bogdan.pavkovic@rt-rk.com)

Dejan Bokan, RT-RK Institute for Computer Based Systems, Novi Sad, Srbija (e-mail: dejan.bokan@rt-rk.com).

AUTOSAR standard daje set specifikacija koje opisuju funkcionalnosti softverskih modula i realizuje zajedničke metode daljeg razvoja na osnovu standardizovanog formata[1]. Arhitektura ovog standarda, odnosno AUTOSAR modela, na najvišem nivou apstrakcije prepoznaje tri različite softverske cjeline[4] (Slika 1.):



Sl. 1. AUTOSAR model [1]

• Osnovni softver (eng. *Basic software module, BSW*)-ovaj sloj se sastoji od modula koji su neophodni za funkcionisanje višeg softverskog sloja. Slojevi od kojih se sastoji osnovni softver su: sloj apstrakcije ECU, složeni upravljački programi, sloj apstrakcije mikrokontrolera (eng. MCAL).

• Izvršno okruženje(eng. Runtime environment, RTE)-realizuje komunikaciju između softverskih komponenti i osnovnog softvera.

•Aplikativni sloj(eng. Application Layer)-funkcionalnost elektronskih kontrolnih jedinica je implementirana u obliku pojedinačnih softverskih komponenti.

A. *Nadzorni časovnik*

Za automobilske sigurnosne sisteme kritično je pitanje zadovoljavanja zahtjeva u realnom vremenu na deterministički način. Da bi se udovoljilo vremenskim ograničenjima, razvijeni su različiti mehanizmi praćenja, kao što su nadzorni hardver ECU jedinice[7], nadgledanje krajnjih rokova[8][9], nadgledanje vremena izvršenja 0. Ovakav vid nadzora kreiran je kako bi se osigurao tačan raspored zadataka0 .

Vertikalu nadzornog časovnika u AUTOSAR slojevitoj arhitekturi čine rukovodilac nadzornog časovnika(nalazi se u servisnom sloju eng. *Service Layer*), sprega nadzornog časovnika(smještena u ECU sloju apstrakcije) i upravljač nadzornog časovnika(smješten u sloju apstrakcije mikrokontrolera)[12] Ovi moduli pružaju usluge za praćenje vremena i ispravnosti izvršenja entiteta u aplikaciji i osnovnom softveru.

B. *Rukovodilac nadzornog časovnika*

Rukovodilac nadzornog časovnika (eng. *watchdog manager*, *WdgM*) je osnovni softverski modul u servisnom nivou koji nadgleda tok programa[13]. Kada se otkrije narušavanje unaprijed definisanih vremenskih ili logičkih ograničenja u programskom toku, potrebno je evidentirati grešku i preći u bezbjedno stanje nakon vremenskog kašnjenja. Sigurno stanje se postiže ponovnim pokretanjem ili izostavljanjem aktiviranja modula nadzornog časovnika.

Po *AUTOSAR* definiciji, tačke u kontroli toka nadgledanog entiteta gdje se aktivnost prijavljuje rukovodiocu nadzornog časovnika su kontrolne tačke.

Polja koja opisuju kontrolnu tačku su:

• *ID* kontrolne tačke
• Lokalni početak, lokalni kraj
• Globalni početak, globalni kraj

Lokalni prelazi predstavljaju prelaze između dvije kontrolne tačke unutar istog nadgledanog entiteta.

Globalni prelazi su prelazi između dvije kontrolne tačke koje pripadaju različitim entitetima.

Nadgledani entitet predstavljen je kontrolnim tačkama kojih može biti jedna ili više. Svaki nadgledani entitet može imati jedno ime i jedno stanje.

Kada se govori o mehanizmima nadgledanja u *WdgM* modulu pominju se tri tipa nadgledanja[14]:

• Nadgledanje u realnom vremenu (eng. *Alive Supervision*) – prati frekvenciju izvršavanja određenog softverskog dijela. To znači da rukovodilac provjerava da li se nadgledani entitet javlja suviše često ili suviše rijetko.

• Nadgledanje krajnjih rokova (eng. *Deadline Supervision*) – nadgleda vrijeme potrebno za izvršavanje nadgledanog entiteta. Glavna svrha je provjera vremenskog, dinamičkog ponašanja entiteta.

• Logički nadzor (eng. *Logical Supervision/Program Flow check*) – nadgleda tok izvršavanja u programu.

Dva su ključna pojma koja treba pomenuti kada je u pitanju nadgledanje i reakcija na greške, a to su vrijeme otkrivanja greške i vrijeme reakcije na grešku.

Vrijeme otkrivanja greške (eng. *Fault Detection*) traje od pojave greške do trenutka kada je ta greška otkrivena i prijavljena sistemu.

Vrijeme reakcije na grešku (eng. *Fault Reaction*) traje od trenutka otkrivanja greške do ponovnog pokretanja sistema. *WdgM* reakcija na grešku:

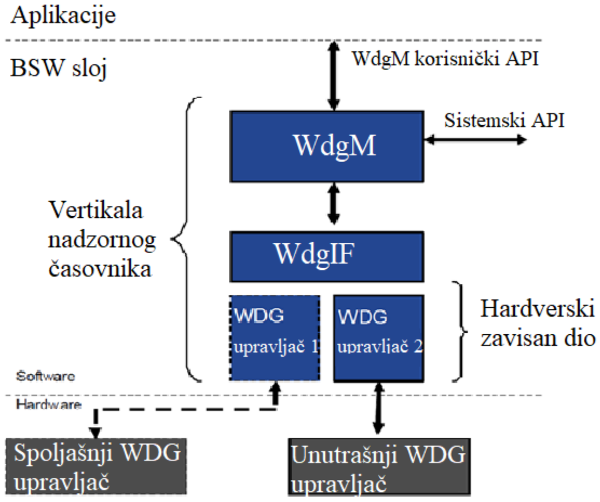
- Obavještenje iz funkcije povratnog poziva
- Ponovno pokretanje sistema
- Stopiranje okidanja nadzornog časovnika

C. Sprega nadzornog časovnika

Sprega nadzornog časovnika (*WdgIf*) je dio ECU apstraktnog sloja. Uvijek se nalazi ispod rukovodioca i iznad upravljača nadzornog časovnika. Sprega komunicira sa upravljačkim programima ispod. Implementacija sprege zavisi od broja upravljača[14].

D. Upravljač nadzornog časovnika

Upravljač je zadužen za pristup samoj periferiji direktno[15] (unutrašnjem i spoljašnjem nadzornom časovniku) i nalazi se u sloju apstrakcije mikrokontrolera. Modul za spoljašnji nadzorni časovnik koristi druge module za pristup spoljnom uređaju.



Sl. 2 Slojevita struktura WDG 1

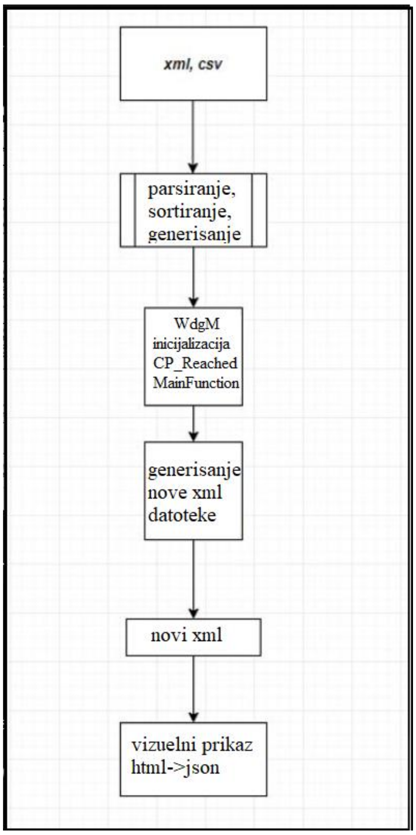
III. KONCEPT RJEŠENJA

Uključen nadzorni časovnik u stvarnom sistemu može stvarati velike probleme prilikom analiziranja nekog problema. Stalno gašenje i ponovno pokretanje jedan su od pokazatelja zašto je to tako. Kako bi se nastavila analiza neke greške na projektima se obično podliježe gašenju nadzornog časovnika i nakon toga se nastavlja sa analiziranjem. Ovaj rad predstavlja jedan pomoćni alat prilikom te analize koji je omogućio da pomoću komandne linije unese željeni poremećaj i isprati lanac događaja koji slijede nakon njega.

Na osnovu već generisanog operativnog sistema odrađeno je parsiranje redoslijeda zadataka i AUTOSAR generisane konfiguracije za stek modul nadzornog časovnika. Parsirani podaci bili su neophodni za grafički prikaz raspoređivača. Grafička predstava kontrolnih tačaka odrađena je tako da se vodilo računa o redoslijedu, kao i koja je kontrolna tačka dodijeljena kom zadatku. Pomenuti grafički prikaz omogućava da se jednostavno uoči poremećaj, tako da je grafički prikaz jedan od osnovnih alata koji su korišteni prilikom analize i testiranja. Ulazni parametri za parsiranje su definisani u csv i arxml formatu (*SE WCET*, *SE period*, *WCET neto*, *WCET abs*).

Sledeći korak jeste simuliranje nadgledanja sistema definisane AUTOSAR arhitekture. Osnovni cilj ovog koraka jeste da prikaže što približnije slijed događaja i grešaka na računaru, kao što je očekivano i u stvarnom sistemu. Glavna razlika je ta što greške koje vidimo na računaru nemaju nikakvu bezbjednosnu posljedicu po izvršenje, već služe isključivo u svrhu analize i rezultiraće ispisima i informativnim porukama, umjesto gašenjem sistema.

Nakon implementacije grafičkog prikaza i simulacije mehanizama, pristupa se analizi sistema, praćenju ponašanja sistema pod uticajima raznih poremećaja koji su namjerno izazvani i koji se odnose na mehanizme nadgledanja. Namjernim izazivanjem grešaka olakšava se analiziranje istih, očekivanih u procesu rada na stvarnom projektu. Data je mogućnost da se predvide različiti lanci događaja, kao i da se smanji vrijeme koje bi bilo potrošeno na pokušaje analize problema uslijed stalnog gašenja sistema.



Sl. 3 Dijagram rješenja

A. Rukovodilac nadzornog sistema u višejezgarnom sistemu

Rukovodilac može biti korišten u jednojezgarnim i višejezgarnim sistemima. U ovom radu obrađen je rukovodilac u višejezgarnom sistemu.

Svaka instanca treba da bude nezavisna jedna od druge i mora biti inicijalizovana njenom sopstvenom konfiguracijom. Poziv *Main* funkcije je odvojen. U stvarnom sistemu se softverske komponente izvršavaju paralelno i vremenski nezavisno. Svako jezgro ima svoje sopstveno vrijeme.

```
- <CORES>
  - <CORE>
      <ID>0</ID>
    - <GENERAL-DATA>
        <NUMBER-OF-TICKS>320</NUMBER-OF-TICKS>
        <MACROTICK unit="us">250</MACROTICK>
      - <TASKS>
        - <TASK>
            <ID>1156</ID>
            <NAME>Task_SchM_NonCritical_C0</NAME>
            <PRIORITY>170</PRIORITY>
            <RANK>1</RANK>
            <PERIOD unit="us">5000</PERIOD>
            <WCET unit="us">400</WCET>
          </TASK>
        - <TASK>
            <ID>1147</ID>
            <NAME>RE_BapFreigabe</NAME>
            <PRIORITY>105</PRIORITY>
            <RANK>1</RANK>
            <PERIOD unit="us">10000</PERIOD>
            <WCET unit="us">100</WCET>
          </TASK>
        - <TASK>
            <ID>1148</ID>
            <NAME>RE_BapTask</NAME>
            <PRIORITY>105</PRIORITY>
            <RANK>2</RANK>
            <PERIOD unit="us">10000</PERIOD>
            <WCET unit="us">100</WCET>
          </TASK>
```

Sl 4. Isječak iz arxml fajla

B. *Sortiranje kontrolnih tačaka prema rasporedu*

Parametri koji su značajni i koji opisuju izvršenje sistema u vremenu su zadati u xml datoteci, a potrebni entiteti su u csv datoteci. Ove dvije ulazne datoteke moraju biti međusobno povezane i predstavljaju jednu cjelinu. Parsiranjem ovih datoteka dobijeni su svi podaci potrebni za simulaciju nadgledanja. Ti parametri su: naziv, perioda, trajanje, vrijeme početka, prioritet i ID. Isječak iz arxml fajla je prikazan na Sl 4.

Nakon izvlačenja pomenutih podataka sve kontrolne tačke sortirane su po vremenu i po prioritetu. Kontrolna tačka koja ima manji prioritet će biti prekinuta ako se u toku njenog trajanja javi neka druga tačka većeg prioriteta.

Nakon sortiranja sve kontrolne tačke kreću sa izvršenjem, baš kao u stvarnom sistemu, istim redoslijedom kako je zahtijevano u ulaznoj datoteci i po prioritetu. Ono što je takođe bilo bitno prikazati jeste vrijeme trajanja koje je oponašano na osnovu ulaznih informacija.

C. *Simulacija nadgledanja*

Nakon uspješno obavljene inicijalizacije, sortiranja i prozivanja kontrolnih tačaka potrebno je simulirati rad prethodno opisanih načina nadgledanja.

Rukovodilac *Main* je sastavni dio izvršenja i on se prozove po zadatom intervalu od 10 milisekundi i tada se vrši provjera nadgledanje u realnom vremenu, nadgledanje krajnjih rokova, logički nadzor.

Vrijeme u sistemu nadzornog časovnika je predstavljeno u tikovima. Potrebno je simulirati vrijeme tako da odgovara vremenu iz stvarnog sistema.

Na osnovu tog vremena se provjeravaju nadgledanja. Provjeru nadgledanja u realnom vremenu treba obaviti tako da se u slučaju da se kontrolna tačka ne javi u očekivanom vremenskom intervalu na konzoli dobijemo ispis o grešci koja se desila.

IV. PROGRAMSKO RJEŠENJE

A. *Parsiranje rasporeda i sortiranje*

Za ulazne podatke iskorišćene su xml i csv datoteke iz stvarnog sistema. Parsiranje je rađeno u programskom jeziku *Python*, svi podaci koji su izvučeni iz tih datoteka su generisani i urađeno je prozivanje funkcije *WdgM_CheckpointReached()*.

Osnovni problem koji se javio prije prozivanja ove funkcije bio je sortiranje kontrolnih tačaka prema rasporedu. Sortiranje je takođe odrađeno u programskom jeziku *Python* i korištene su funkcije:

•*expiry_points()* - funkcija koja sortira podatke koji su izvučeni iz ulaznih datoteka parsiranjem. Sortira kontrolne tačke po vremenu njihovog javljanja i po jezgrima.

•*priority_sort()* - Prethodno sortirana lista po vremenu javljanja se sortira i po prioritetima . Ako se desi da dvije kontrolne tačke počinju istovremeno prednost će imati tačka sa većim prioritetom, tačka manjeg prioriteta ostaje da čeka svoje red. Entitet može biti prekinut i u toku izvršenja, ako se desi da je došlo do javljanja izvršioca sa većim prioritetom, trenutni entitet ostaje u stanju čekanja sve dok mu se ne signalizira da je prioritetniji entitet završio sa radom.

Nakon sortiranja je generisano kojim se redoslijedom vrši pozivanje *WdgM_CheckpointReached()* funkcije.

B. *Implementacija rukovodioca*

Prvi korak koji je odrađen jeste postupak inicijalizacije. Svaki zadatak inicijalizovan je pomoću funkcije *WdgM_Init()*.

U stvarnom sistemu WDG zadatak ponavlja se kružno na svakih 10 milisekundi. To znači da se poziv funkcije *WdgM_MainFunction()* ponavlja svakih 10 milisekundi. Ova funkcija ima ključnu ulogu jer se u njoj vrše provjere ispravnosti. Trajanje jednog ciklusa naziva se hiper period, Na osnovu trenutnih ulaznih parametara koji su obrađeni u ovom primjeru koji će biti opisan hiper period je 80 milisekundi i nakon toga se završava jedan ciklus nadzora. Nakon izvršenja *WdgM_MainFunction()* očekuje se neka od reakcija rukovodioca.

- Ako dođe do greške u nadgledanju u realnom vremenu greška će biti detektovana na kraju nadgledanog referentnog ciklusa (eng. *Alive supervision reference cycle*).
- U slučaju nadgledanja programskog toka ako dođe do greške ona će biti detektovana na kraju svakog nadgledanog ciklusa.
- Ako je greška u nadgledanju krajnjih rokova ona će biti detektovana na kraju svakog nadgledanog ciklusa, nastavak kršenja ovog vida nadgledanja detektuje se na kraju svakog krajnji rok nadgledanog entiteta.

Ponašanje sistema nakon uočavanja neke od pomenutih grešaka zavisi od konfiguracije i tipa poremećaja.

V. PROVJERA ISPRAVNOSTI

A. *Opis testiranja*

U svrhu testiranja korišteni su ulazni parametri sa stvarnog sistema. Ovakav pristup omogućio je poređenje sa stvarnim sistemskim greškama i utvrditi ispravnost samog rada.

Kako je stvarni sistem čiji si ulazni parametri iskorišćeni sadržao 3 jezgra, a ona u sistemu rade u paraleli. U ovom rade sva tri jezgra su testirana istovremeno i softverski spojena u jednu cjelinu, prikazan je njihov paralelizam. Na računaru se pomoću komandne linije prati ispis rezultata, kao rezultat testiranja dobijaju se poruke o prekršajima.

Prilikom pokretanja radi se inicijalizacija sistema.

Korisnik treba da odabere jezgro na kome će nanijeti poremećaj kao i vrstu poremećaja koju želi, promjena *WCET* vremena ili greška u nadgledanju u realnom vremenu.

Ako je u sistemu sve prošlo bez greške prilikom izvršavanja rukovodioca prozvaće se *TriggerWindow* funkcija na osnovu čega je testirana ispravnost. Vizualni prikaz sistema bez greške dat je na Sl.5

B. Poremećaj nastao promjenom WCET vremena

Najgore vrijeme izvršenja predstavlja ukupno vrijeme koje je dato jednom zadatku da se izvrši. Vrijednost ovog parametra predstavljena je sa dva termina *bruto* i *neto WCET*. Kada je riječ o ukupnom *bruto* vremenu možemo reći da je to vrijeme koje protekne od početka do kraja zadatka sa uračunatim svim prekidima od strane prioritetnih zadataka. Dok je *neto WCET* vrijeme od početka do kraja ali predstavlja samo sabrane vremenske trenutke u kojima je aktivan dati zadatak.

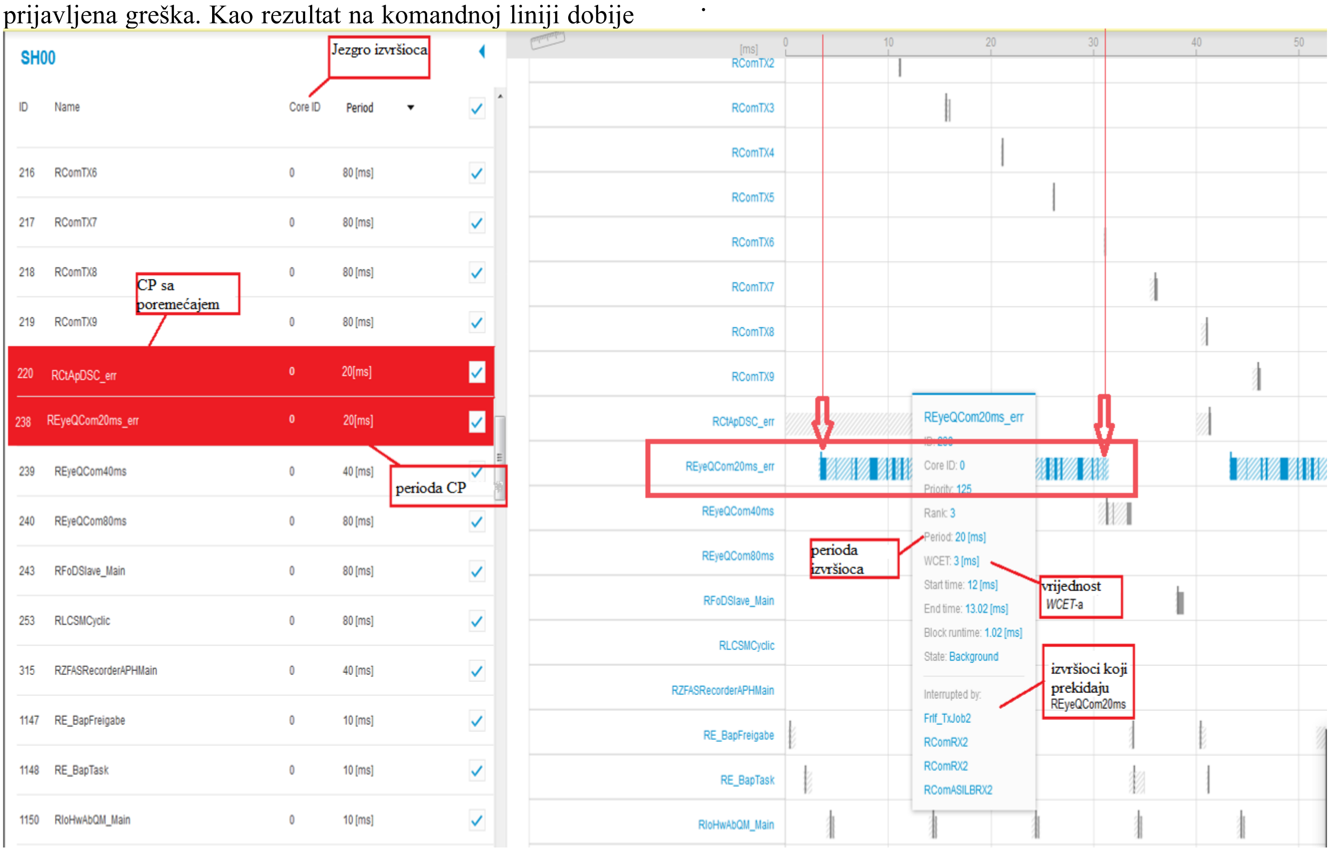
Vrijednost *WCET* vremena koja je konfigurisana za određeni nadgledani entitet i data u rasporedu očekivana je vrijednost u sistemu sa kojom svi nadzori rade ispravno. Nakon što se napravi neka promjena *WCET* vremena i ispisivanja poruka o ispravnosti sistema generiše se nova datoteka *arxml* koja predstavlja ulazni parametar za vizualizaciju sistema. Ako je neka promjena unesena na grafičkom prikazu promijenjena kontrolna tačka mijenja boju sto se vidi na Sl 5. Greška u sistemu vidljiva je na slici gdje su crvenom bojom markirani izvršioci kod kojih je prijavljena greška. Kao rezultat na komandnoj liniji dobije se poruka o svim prekršajima koje je izazvala promjena *WCET* vremena na željenom entitetu.

Na osnovu vizualizacije omogućeno je lakše praćenje dešavanja u sistemu, način na koji se nakon bilo koje promjene izmiješaju kontrolne tačke. Simulirana je vremenska osa i kontrolne tačke u vremenu.

C. Poremećaj u nadgledanju realnog vremena

Kada se nanese ovaj vid poremećaja u sistemu dolazi do situacije u kojoj se određena kontrolna tačka ne prozove. Tada sistem detektuje grešku u nadgledanju u realnom vremenu. Primjer iz tabele takođe pokazuje grešku nad *RCtApDSC* koja ima period 20 milisekundi. Možemo uočiti da je došlo do problema kada je rukovodilac prepoznao da se u prvih 20 milisekundi nije javila ova kontrolna tačka, a mehanizam nadzora u realnom vremenu očekivao je da će doći do njenog javljanja. Nakon otkrivanja problema brojač se nije uvećao i kao status vraćena je vrijednost "nije uspjelo"(eng. FAILED).

Kada se poveća WCET u ovom slučaju nad nadgledanim entitetom pod nazivom *ReyeQCom20ms* koji je prikazan u tabeli (T 1.) desi se poremećaj u prvih 10 milisekundi. Očekivana vrijednost je 2.7 milisekundi jer je ovo nadgledani entitet koji je niskog prioriteta i isprekidan je od strane ostalih koji imaju veći prioritet. Prilikom testiranja promijenili smo vrijednost na 3 milisekunde. Zbog poremećaja na jednoj kontrolnoj tački i greške u nadgledanju krajnjeg roka u prvih 10 milisekundi očita se greška, ali se prozove i funkcija *TriggerWindow*.



Sl. 5 Greške na jezgru 0 nakon promjene *WCET* vremena nadgledanog entiteta *ReyeQCom20ms*

Na Sl 5 može se uočiti kako će *ReyeQCom20ms* zadatak imati uticaj na sistem kada se njemu nanese vremenski poremećaj. Takođe primjetna je lista izvršilaca koji prekidaju pomenuti entitet zbog većeg prioriteta, pri čemu će svaki od pomenutih prekinuti izvršenje. trenutnog. Puna linija predstavlja izvršenje *ReyeQCom20ms* dok je isprekidanim poljima predstavljen period kada je posmatrani entitet u pozadini i čeka na izvršenje zadatka sa većim prioritetom.

jezgra / poremećaj	JEZGRO 0	JEZGRO 1	JEZGRO 2
sistem bez poremećaja	TriggerWindow (za svaki zadatak koji pripada tom jezgru)	TriggerWindow (za svaki zadatak koji pripada tom jezgru)	TriggerWindow (za svaki zadatak koji pripada tom jezgru)
greška na kontrolnoj tački ReyeQCom 20ms WCET=3ms	*ReyeQCom20ms Execution flow violation*	TriggerWindow (za svaki zadatak koji pripada tom jezgru)	TriggerWindow (za svaki zadatak koji pripada tom jezgru)
	TriggerWindow	TriggerWindow (za svaki zadatak koji pripada tom jezgru)	TriggerWindow (za svaki zadatak koji pripada tom jezgru)
	RBAQM10ms Execution flow ERROR	TriggerWindow (za svaki zadatak koji pripada tom jezgru)	TriggerWindow (za svaki zadatak koji pripada tom jezgru)
	SetResetReasone		
Alive monitoring greska nad RCtApDSC	*RCtApDSC Execution flow ERROR*	TriggerWindow (za svaki zadatak koji pripada tom jezgru)	TriggerWindow (za svaki zadatak koji pripada tom jezgru)
	SetResetReasone		

T 1. Ponašanje sistema nakon poremećaja 1

VI. ZAKLJUČAK

U okviru ovog rada prikazano je rješenje i način praćenja poremećaja prijavljenih od strane nadzornog časovnika. Česte su situacije da se u radu na realnoj platformi nailazi na poteškoće po pitanju očekivanog ponašanja hardvera na određene zahtjeve iz softvera. Kada govorimo o samom nadzornom časovniku i reakciji fizičkog upravljača nekada sa sigurnošću ne možemo da tvrdimo šta je uzrok okidanja greške i gašenja sistema. Sigurnosno gašenje može značajno usporiti proces analiziranja nekog problema koji sam po sebi ne mora biti vezan isključivo za nadzorni časovnik. Takav vid poteškoća je moguće pratiti samo isključenjem upravljača. Simulacijom je omogućeno da se minimalizuju ta odstupanja, predvide greške u sistemu i olakša analiza sistema. Greška koja se desi na jednoj kontrolnoj tački može da prijavi grešku tek na sledećoj kontrolnoj tački koja je u redu. Rad omogućava da se isprati i predvidi takav vid prekršaja.

Kao što je prethodno pomenuto za ovaj rad je korišten već postojeći raspored zadataka koji sam po sebi predstavlja preduslov za početak simulacije. Budući rad obuhvatiće unaprjeđenje postojećeg rješenja simulacijom operativnog sistema, gdje će se voditi računa i o simulaciji raspoređivanja zadataka kako bi se poremećaj mogao nanijeti direktno u raspoređivanju i ispratiti cijeli proces. Ovaj vid unaprjeđenja značajno bi mogao poboljšati analizu i predviđanje grešaka.

LITERATURA

[1] AUTOSAR, "*Layered Software Architecture*," *[Online]*, https://www.autosar.org/fileadmin/user_upload/standards/classic/4-3/AUTOSAR_EXP_LayeredSoftwareArchitecture.pdf., 2017-12-08 [Accessed July 2021]

[2] B. Catalin-Virgil, F. Ioan and F. Heininger, "*A new trend in automotive software: AUTOSAR concept*" SACI, IEEE 8th International Symposium, Timisoara, Romania, May 2013.

[3] Techemergence, "*The Self-Driving Car Timeline –Predictions from the Top 11 Global Automakers*", 2018.

[4] H. Fennel et al. "*Achievements and exploitation of the AUTOSAR development partnership*" SAE Technical Paper Series 2006-21-0019, SAE International, October 2006,[Accessed July 2021]

[5] Mazen Ahmed, Mona Safar, "*Formal Verification of AUTOSAR Watchdog Manager Module Using Symbolic Execution*", IEEE 30th International Conference on Microelectronics, 2018

[6] Mazen Ahmed, Mona Safar , *"Symbolic Execution based Verification of Compliance with the ISO 26262 Functional Safety Standard"*, IEEE 14th International Conference on Design & Technology of Integrated Systems In Nanoscale Era, 2019

[7] J. Ganssle, "*Watching the Watchdog*", Embedded World, 2003.

[8] *AUTOSAR* , *"Specification of Communication" [Online],* https://www.autosar.org/fileadmin/user_upload/standards/classic/3-2/AUTOSAR_SWS_COM.pdf [Accessed July 2021]

[9] Michael Kunz, "*OSEK OS"*, March 18, 2009, http://www.uni-obuda.hu/users/schuster.gyorgy/rtos/OSEK.pdf [Accessed July 2021]

[10] The AUTOSAR Consortium, "*AUTOSAR Specification of Operating System"*, pp. 33-35, 2006.

[11] Nahmsuk Oh, P. Shirvani, E. McCluskey, "C*ontrol-Flow Checking by Software Signatures*", IEEE Transaction on Reliability, vol. 51, Mar-2002

[12] Texas Instruments MCUSW, "*Wdg Design Document " [Online],* *http://software-dl.ti.com/jacinto7/esd/processor-sdk-rtos-jacinto7/latest/exports/docs/mcusw/mcal_drv/docs/drv_docs/design_wdg_top.html* [Accessed July 2021]

[13] *AUTOSAR*, "*AUTOSAR SWS Watchdog Manager*" *[Online],* https://www.autosar.org/fileadmin/user_upload/standards/classic/4-3/AUTOSAR_SWS_WatchdogManager.pdf , *2017-12-08* [Accessed July 2021]

[14] *AUTOSAR, " Specification of Watchdog Interface"* https://www.autosar.org/fileadmin/user_upload/standards/classic/4-3/AUTOSAR_SWS_WatchdogInterface.pdf, [Accessed July 2021]

[15] *AUTOSAR, "Specification of Watchdog Driver "[Online]* https://www.autosar.org/fileadmin/user_upload/standards/classic/4-3/AUTOSAR_SWS_WatchdogDriver.pdf , [Accessed July 2021]

ABSTRACT

With the development of the automotive industry and software within it, as a technical consequence, there was a need for mandatory integration of protection mechanisms in the embedded cores of the operating system. One of the basic mechanisms for system protection is the watchdog. This component aims to monitor all other components initiated by the scheduler and thus enable the safe operation of the system. As the problems reported by Watchdog are relatively difficult to track and analyze in a real system, the idea came up to simulate the operation of the Watchdog component on a computer with the same input parameters as in a living system. This paper provides a solution for simulating the system monitoring mechanisms of the defined AUTOSAR architecture. The simulation makes it possible to minimize deviations, predict errors in the system and facilitate the analysis itself. Work can contribute to faster system development because it allows to predict errors that will occur in the system before implementation.

ONE SOLUTION FOR ANALYZING AND DISPLYING CHECKPOINTS IN THE AUTOSAR WATCHDOG CONFIGURATION

Ivana Tešević, Branko Milošević Dejan Bokan, Bogdan Pavković

Implementation of Smooth Streaming protocol through a generalized software framework

Miroslav Suša , Ilija Bašičević, *Senior Member, IEEE*

Abstract—Adaptive streaming is a technology for transmitting multimedia content over a network such as the Internet. This way the content is available at any time which has brought big changes. One of the many streaming technologies is Smooth Streaming. In addition to the transmission of content via one of the protocols, it is necessary to ensure its reproduction. In this paper, the implementation of the Smooth Streaming protocol within a single media player is presented. The implementation was performed through a generalized software framework, which will also be discussed. The role of the framework is to facilitate the integration of the remaining adaptive streaming protocols into the media player.

***Index Term*—adaptive streaming; content playback; Smooth Streaming;**

I. Introduction

The fourth industrial revolution brought many changes in terms of consuming multimedia content. When it comes to content transfer, the most important innovation is Streaming technology. Streaming makes it easier for users to access content whenever and at any time they want, which previous technologies could not provide. In the field of television, new technology is completely taking over the market from the traditional way of content broadcasting[1].

As a consequence of the development of a new way of data transfer, standards have emerged according to which this transfer will be performed. Some of the best known and most prevalent today are the MPEG-DASH and Smooth Streaming standards. In addition to the transmission of the content itself, it is necessary to ensure its reproduction.

This paper presents an implementation of the Smooth Streaming protocol, the creation of a generalized software framework for managing adaptive streaming protocols, as well as the integration of the software framework with the Smooth Streaming protocol and a media player for content playback.

The rest of the paper is organized in the following order: Section II discusses streaming technology and its types. Section III shows the architecture of the media player on which the work solution is implemented. Section IV defines a programmatic framework for managing adaptive streaming protocols. Section V discusses creating a library for the Smooth Streaming protocol. Section VI shows the architecture of the media player after applying the solution. Section VII deals with the validation of the solution and the results. Section VIII provides a conclusion on the work.

Miroslav Suša – RT-RK Institute for Computer Based Systems, Novi Sad, Serbia (e-mail: miroslav.susa@ rt-rk.com).

Ilija Bašičević – Faculty of Technical Sciences, University of Novi Sad, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia (e-mail: ilibas@uns.ac.rs).

II. Streaming

Streaming is a technique of continuous transmission of video and audio material via wired or wireless internet connection. Before the advent of streaming, playback of content from the Internet was possible in two ways. The first way is to upload the complete file to the device and only then is playback possible. Another way is to use a progressive download.

A. Progressive streaming

Progressive download allows you to play content while downloading it to your device [2]. Downloading is done regularly, which means that not only the selected part of the file can be downloaded, but the complete one. Any part of the downloaded content can be played as desired. The content transmitted in this way is of fixed quality and resolution. In other words, only one video file can be uploaded. Since different content resolutions require better or worse internet traffic, in a situation where the flow is poor, progressive downloads will often lead to transmission interruptions. In addition, the content will be displayed differently on different devices.

Higher resolution files take up more memory space, and the file transfer speed depends on the internet flow which tells us how much data the user can receive in a unit of time. If the flow is poor and the video has a higher resolution, part of it will not be able to be transmitted in its entirety and playback will be delayed. In addition to downtime, progressive downloads also cause the problem of presenting videos on devices with different screen resolutions. For example, when playing a video that is 720p resolution, on a screen with 1080p resolution, the image will be stretched and pixelated.

B. Adaptive streaming

Adaptive streaming is a streaming technology based on the HTTP (HyperText Transfer Protocol) protocol. The benefit of using HTTP technology is the unhindered passage through the firewall and NAT (Network Address Translation) devices that remap IP (Internet Protocol) addresses. In addition, the complete implementation of HTTP logic is on the side of the content seeker, which reduces the need for a continuous connection between the provider and the service provider.

Adaptive streaming, instead of a complete file, transmits and plays its parts for a few seconds [3]. We call such parts of a file its segments. Since the content is divided into segments, any part of it can be added as desired. Just before the segment expires, the next one to be played is delivered. After moving on to the next segment, the previous one is deleted, and the process takes place until the complete content expires. This gives the impression of continuous playback of content. Information about all video and audio files and their segments, as well as details of the need for their transfer and playback can be found in the manifest file.

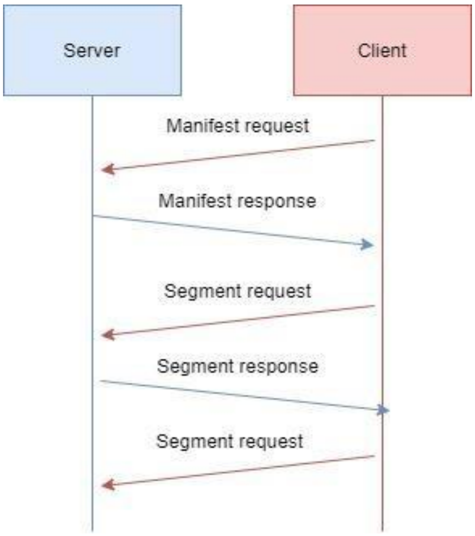


Fig. 1. File exchange between client and server during adaptive streaming

Adaptive streaming solves the problems of progressive download [4]. Files of different resolutions are created for the same content. Depending on the size of the client's screen, a file segment of the appropriate resolution is provided.

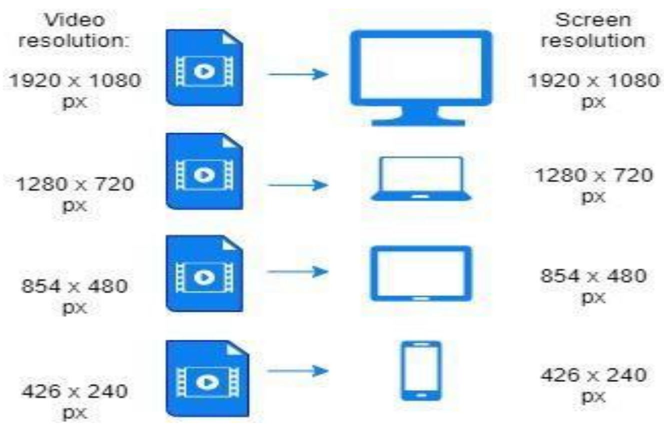


Fig. 2. Display video resolution selection for different screen resolutions.

In the typical communication scenario between a client and a server using adaptive streaming the client first sends a request for a manifest and receives it in response from the server. The client then sends requests for fragments which the server delivers.

III. Media player arhitecture before protocol implementation

IWedia Player (IWP) is a library written in the C ++ programming language that aims to provide a high-level player interface. The player allows you to play audio and video content. It is used as a part of an application written for the Android platform and provides it with a user interface.

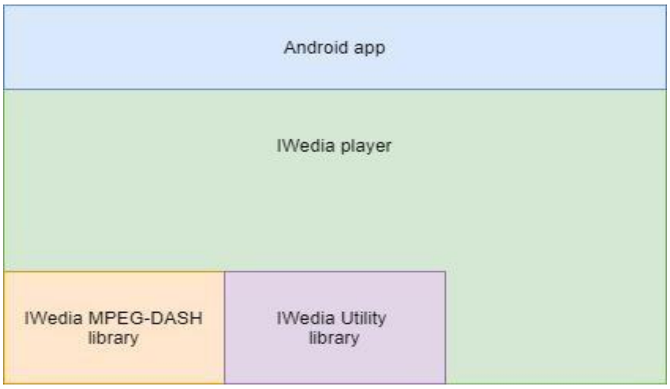


Fig. 3. Display of player architecture before implementing the solution.

In order for the player to enable the MPEG-DASH streaming protocol, an IWP-DASH library was created. In addition to defining the elements of this streaming protocol within the dash library, the logic for adaptive streaming has been implemented, which is closely related to the dash protocol. Components shared between other IWedia libraries such as the player and dash are housed in the IWedia utils library.

IV. Software framework for managing adaptive streaming protocols

When it became necessary for the media player to support, in addition to the MPEG-DASH protocol other adaptive streaming protocols as well, the implementation logic had to be generalized and displaced from the dash library.

The goal of the adaptive streaming protocol management framework is to provide the media player with an interface through which to obtain content for playback. All protocols aim to transfer content and regardless of their complexity, we can see numerous similarities between them. Since MPEG-DASH is an official international standard, it is the most developed and provides the most opportunities during implementation [5]. All other standards can be viewed as its subset.

The software framework can be divided into four logical units that have a role in downloading manifests, segments, adapting to network conditions and creating an entrance to the library.

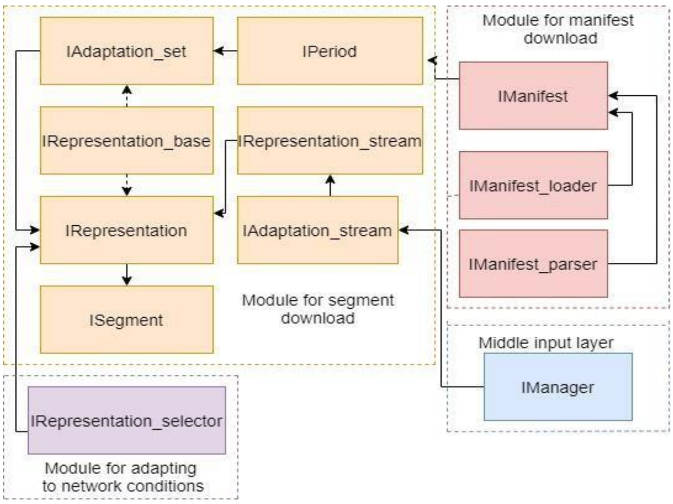


Fig. 4. Generalized software framework solution architecture.

A. Module for manifest download

Manifest, as mentioned, is the central document for gathering information on the content to be transmitted. It is available on the server along with the provided content. The type of manifest can be dynamic, if we broadcast live content, or static, if the content is downloaded on demand. To stream live content, the manifest needs to be delivered periodically because the content is constantly changing.

To provide a manifest, URI (Uniform Resource Identifier) from which it can be downloaded is required. The download is performed with the help of a previously implemented download class, which needs to be provided with data about the speed, number of attempts and download time of the manifest. As a result of successful delivery, a string with the contents of the manifest is obtained. After it is downloaded, the manifest is parsed.

At the level of libraries that implement the standard, it is necessary to implement interfaces that represent the manifest and the factory for creating the manifest.

B. Module for segment download

Within period elements of the manifest, that contain the initial time and duration of the content, there are adaptation elements. Their primary purpose is to provide information about the type of stream being transmitted. Inside the adaptation element are elements that represent the stream. They contain data on the flow rate required to download the stream in a certain quality as well as information on the segments that need to be downloaded. At the stream type level, a structure is created that will download the segments.

Libraries of specific protocols that implement the created interfaces define the form of stream representation as well as the form of segments. The representation form creates segments based on a given start time, carries information about the number of available segments, as well as the broadcast time period provided by the representation.

Segment download control is defined in this software framework. The time period in which the download will be performed is determined, the representation within which the segments will be downloaded is selected, the ordinal number of the next segment to be created is calculated and the creation of the segment is initiated.

C. Module for adapting to network conditions

The network adaptation module is key to performing adaptive streaming. Depending on the speed of the user's Internet flow, it is necessary to correct the representation of the stream being downloaded. The factor that influences the choice of representation, in addition to the flow rate, is the type of content that is downloaded. For the purpose of selecting a representation, a selector is created that stores all available representations and, based on the current flow and type of content, selects one of them to be played. After the download, the number of bits downloaded, as well as the time period required for the download, are forwarded to the flow rate meter. Based on the obtained parameters, the meter calculates the current flow rate that is available when initializing the next download.

Middle input layer

As part of the software framework, a manager has been created to manage the processes. Its presence is necessary in the media player that uses the library. The task of the manager is to initiate the loading of the manifest when it comes to streaming on demand or perform periodic loading of the manifest if a live broadcast is performed, as well as the interruption of these operations. In addition, the manager creates a program representation of the adaptation stream that has the ability to further manipulate stream representations and segments.

D. Software framework integration with the media player

In order to enable the reproduction of content by adaptive streaming protocols, it is necessary to integrate the software framework with media player. The integration is done by adding a component that has access to the software framework manager. In this way, the processes realized by the software framework are initiated, such as taking over the manifest and adding segments. The ultimate goal of process initiation is to obtain segments and prepare them for reproduction.

Within the media player, there is also logic for determining the adaptive streaming protocol that will be used, as well as factories that will, depending on the selected protocol, create a component that has access to the software framework.

V. Smooth Streaming library

The Smooth Streaming library is a C ++ implementation of the Microsoft Smooth Streaming protocol used by the IWedia player to play content. The library consists of "*ismc*" and *"abr"* modules. The *Ismc* part of the library was named after the extension of the Smooth Streaming protocol manifest client. Within this part, the manifest is parsed and elements and attributes representing the data collected by parsing are realized. *Abr* part of the library represents the implementation of a software framework for managing adaptive streaming protocols.

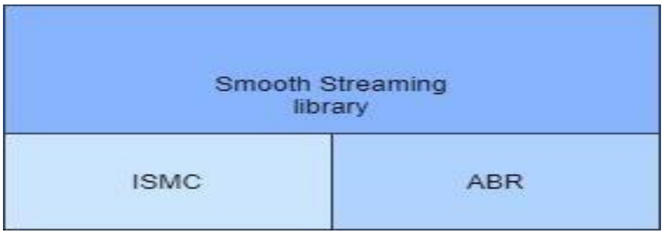


Fig. 5. Smooth Streaming Library Components.

A. Library creation

In order to implement the protocol, it is necessary to parse the protocol manifest and present its elements within the library. In addition, it is necessary to provide the types of content exchange messages defined by this transport protocol, which are: manifest request, manifest response, segment request and segment response.

1) Manifest request

A manifest request is sent to obtain a manifest containing all the necessary information to reproduce the content. In order to send this request, a URI to the manifest is required as well as information on which extension of the manifest file is expected. Manifest extensions differ in whether it is a server manifest that has an *ism* extension or a client manifest whose extension is *ismc*.

2) Manifest response

The manifest response is obtained in the form of an *ismc* file with metadata related to the playback of the content. The file is a well-formed XML (Extensible Markup Language) and consists of the following elements: SmoothStreamingMedia, Protection, StreamIndex, QualityLevel and StreamFragment. All of the above elements are presented within the library as classes, and their correlations are clearly visible and described below.

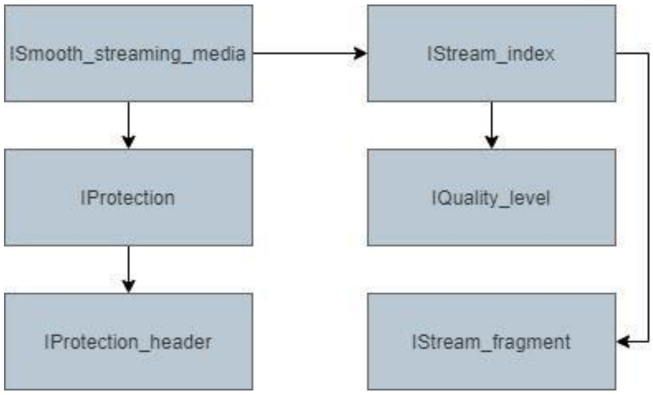


Fig. 6. Smooth Streaming library solution architecture.

a) SmoothStreamingMedia

SmoothStreamingMedia is a root element that contains all the other elements of the manifest. The direct descendants of this element are the StreamIndex and Protection elements. Its attributes carry information about the main and secondary versions of the manifest as well as whether the manifest describes live or on-demand content. Within the attribute, the duration of the content described in the manifest is also defined.

SmoothStreamingMedia is implemented within the library so that its creation requires URI of manifest as well as xml files in string format. The string is then parsed using the sub-element names and attributes listed as constants.

a) Protection

Protection is an xml element that includes the metadata needed to play protected content. It contains information on the unique identification of the security system used on the given content, as well as the encoded data that the system uses to enable the reproduction of the content to authorized users.

b) StreamIndex

StreamIndex is the most important element within a manifest because it contains metadata for playing a specific stream. This means that the element provides information about the type of content that is transmitted by a particular stream, that is, whether it is an audio, video or text stream. Based on the stream type, the availability of attributes within an element changes. Only in the case of video, there is information about the maximum available content resolution that is available, as well as the recommended playback resolution. The number of qualities, segments, as well as the duration of the stream are available within this element.

b) QualityLevel

The QualityLevel element carries metadata about the playback of a specific track within the stream. Depending on the type of stream in which it is located, its attributes differ. For video within the video stream, the required resolution attributes as well as parameters specific to a particular media format are required. When it comes to audio recording within the audio stream, in addition to the previously mentioned attributes, we also have data on the number of channels of the audio tape, sampling rate, sample size, limits for optimizing audio decoding and identification of media format used. The attributes that each record contains are those that carry information about the unique identification of each record and the download speed required to retrieve a particular record.

c) StreamFragment

The StreamFragment element contains metadata about a set of related segments in the stream. Its attributes carry information about the start time of the segment, its duration, the order in a series of segments as well as the possibility of repetition. For a segment to be valid, it must contain either a duration attribute or a start time attribute. A series of segments is called adjacent if the start time of any segment, with the exception of the first, is equal to the sum of the start time and the duration of its predecessor.

3) Segment request

A segment request is created to retrieve the desired segment from the server. To create it, it takes URI to the desired segment, its bitrate, the name of the stream within which the fragment is located, the start time of the desired stream, as well as the type of response that the client expects from the server.

4) Segment response

A segment response is a response that is received after sending a request to obtain a segment. The answer can be complete or partial. If the answer is complete it contains media and segment metadata, while partial responses contain only media or metadata.

B. Framework implementation

In the Smooth streaming library it is necessary to implement two of the four modules of the framework and they are: Module for manifest download and Module for segment download.

2) Implementation of the module for manifest download

As mentioned earlier, it is necessary to implement the manifest factory interface as well as the manifest interface.

a) IManifest interface

IManifest methods gather the necessary information that each manifest should have, namely: whether the manifest is live or on-demand, the duration of the manifest, the minimum time required to load the manifest, as well as adding the manifesto period. All data can only be obtained by parsing the manifest.

b) IManifest_factory

The manifest creation factory contains only one method that instantiates a class that implements the IManifest interface. It forwards manifest uri and the contents of the manifest that is necessary to parse.

3) Implementation of the module for segment download

Module for segment download defines the necessary logic to supply the parsed element data, as well as the logic for creating segments.

a) IRepresentation

Interface methods obtain, from the QualityLevel element, data described in the part of the paper with the same name. All data is present in the node and is very easy to obtain.

b) IAdaptation_set

The IAdaptation_set interface is composed from set of methods that retrieve data from the Stream_index element of the library. All methods return the present attributes or sub-elements of a given element.

c) ISegment

This interface is defined by a set of methods for retrieving the Stream_fragment element attribute with the exception of the get_uri method. The get uri method calculates the uri to a given segment that is different from the manifest uri

d) IRepresentation.

The role of the IRepresentation interface is to create segments, add the total number of segments, add the duration of all segments and find the segment with a given index

The number of segments is obtained when initializing the class of this interface by going through all segments and taking into account their repeat attribute which tells how many times a given segment is repeated.

During the process of calculating segments, the total duration of all segments can be easily obtained. The timestamp of the first and last segment is taken, or their length and repeat tag if the timestamp is not available.

A segment with a given index is supplied by going through all available segments, taking their duration and repeat attribute, calculating the index of each segment and returning the resulting one. The limitation of this method is to pass an index that is not less than zero and that is not greater than the total number of segments.

VI. MEDIA PLAYER ARHITECTURE AFTER PROTOCOL IMPLEMENTATION

By removing the definition of adaptive streaming protocol from the dash library and generalizing it, a software framework for managing adaptive streaming protocols is obtained. This makes it easier to use and add new adaptive streaming protocols such as the Smooth Streaming protocol. In addition, their integration with the player is facilitated.

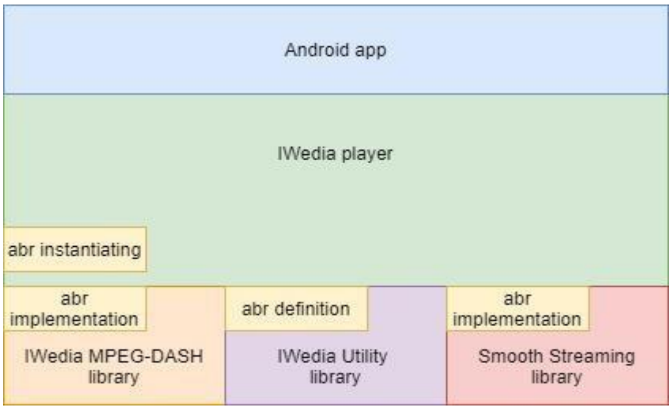


Fig. 7. Display of player architecture after solution implementation.

VII. TESTING

A. Description of the test environment

The environment for testing of this solution comprises Smooth Streaming content that can be accessed via the network, an application for playing content, as well as an Android P development board on which the application will be launched.

The content playback application is written in the Java programming language for the Android platform. It provides a simple user interface from which content playback can be controlled, and uses the IWedia player library interface for playback itself.

Content preparation, implementation of the Smooth Streaming library and software framework, as well as their integration with the media player are described in the previous chapters. The Android P development board connects to the same network from which the prepared content is available to it, and the Android application is installed and launched on it. With this step, the test environment is ready and testing can begin.

B. Testing procedure

After installing the Android application on the board and launching it, you get access to the list of all available streams. Clicking on the desired stream starts playback. Playback can be interrupted, paused or restarted at any time.

C. Test results

As stated, by clicking on the desired stream, in this case on the stream belonging to the group of Smooth Streaming streams, playback starts. The start of content playback always takes place at a lower resolution until the user's internet flow is determined. After that, if it is determined that the conditions are met, playback continues at higher resolutions, which tells us that the SmoothStreaming protocol has been successfully implemented.

Selecting one of the available streams that are transmitted by other adaptive streaming protocols results in content playback. Successful playback start shows that the generalized software framework has been correctly implemented.

VIII. CONCLUSION

Within this paper, a solution for integration of Smooth Streaming standards for broadcasting content is described. Also, a generalized program framework has been implemented, which enables easier integration of the remaining standards.

The protocol library and software framework are written in C++. In this way, speed and flexibility are achieved. Like the media player in which they are implemented, they can be used on both Android and Linux platforms.

The integration framework in the media player enables easier integration of existing, as well as the protocols that may arise in the future which can be seen as a subset of the MPEG-DASH protocol.

ACKNOWLEDGMENT

On this occasion, I would like to thank my colleague Nikola Špirić on the provided support.

REFERENCES

[1] Nenad Lovcevic, Jelena Simic, Miroslav Dimitraskovic and Ilija Basicevic “Modul za prijem i obradu JSON komandi u programskoj podršci digitalnog TV prijemnika” 61st IcEtran conference in Kladovo, Serbia, June 5 – 8, 2017.

[2] Stefan Lederer, Christopher Mueller,Christian Timmerer, Hermann Hellwagner “Adaptive Multimedia Streaming in Information-Centric Networks” IEEE Network, November 2014.

[3] Ilija Bašičević, Nenad Lovčević, Nenad Šoškić, Milan Ačanski “Internet as Infrastructure for Digital Television” 62st IcEtran conference in Palić, Serbia, June 11 – 14, 2018.

[4] Rubem Pereira, Ella Pereira “Dynamic Adaptive Streaming over HTTP and Progressive Download: Comparative Considerations”, IEEE 28th International Conference on Advanced Information Networking and Applications Workshops (WAINA) - Victoria, Canada, May 1 2014

[5] Sunho Seo; Younghwan Shin; Jusik Yun; Wonsik Yang; Jong-Moon Chung “Adaptive high-resolution image transmission method using MPEG-DASH” International Conference on Information and Communication Technology Convergence (ICTC) Jeju, Korea (South), October 18-20, 2017.

Implementation of the GDPR Compliant Data Handling for Smart Home Solution

Sandra Bugarin, Sandra Ivanović, Marija Antić

***Abstract*—The amount of personal data collected and shared in the Internet of Things (IoT) is causing increasing concerns regarding the user privacy in IoT. The recently introduced General Data Protection Regulation (GDPR) is a legal framework that sets guidelines for the collection and processing of personal information and aims to strengthen user rights. In order to comply with the GDPR requirements, the existing smart home system is extended with the cloud service, responsible for user consent management and appropriate data handling. The architecture of the solution, as well as the results of functional and performance testing are presented in this paper.**

***Index Terms*— GDPR; smart home automation; IoT**

I. Introduction

The users surfing the web are under a risk of privacy violation, as the websites are collecting data about them and may be sharing it with third party services. Recently, the General Data Protection Regulation (GDPR) has entered into force, with the aim to protect the user privacy, and allow them better control over the collected data and the scenarios it is used in [1]. According to GDPR, the services are required to inform the users about the types of data collected and the purpose of this action, so the users can choose to engage only with websites and services that do not violate their privacy, or opt out of the use of their information for particular purposes.

While the data collected by the websites usually serves only marketing purposes, and is not necessary for the normal operation of the website, the problem of GDPR compliance in Internet of Things (IoT) solutions is of a more complex nature [2]. Namely, IoT systems typically connect multiple devices owned by a single user, and allow them to perform a certain function together. Therefore, the exchange of data is in the essence of IoT. On the other hand, there exists a tendency in the IoT solutions to collect more data than actually needed for the normal system operation, as it may become useful in the future scenarios [3], [4]. This data should be carefully stored and protected, as well as anonymized [5], and the users should be provided with the mechanisms to inspect or delete the collected data at any time [6]. Also, it is necessary to be transparent about the ways data is processed, to inform the users timely when the privacy policies change, and to allow them to opt out of the service if they do not agree to the changes.

Studies have been conducted that show that the user attitude towards data collection depends on multiple factors, such as the environment the data is related to (home, office, traffic), types of data collected (video, photo, sensory data, voice), who has access to it (government, businesses), as well as the purpose of data collection (safety, convenience, marketing) [7]. Smart home users are willing to allow data collection as long as it is used only within the system, for the purpose of connectivity and convenience [8], but seem not aware of the possible privacy issues associated with machine learning and potentially sensitive information that can be revealed by data analytics [9]. This information should be communicated through the privacy policy and terms of use, in a manner that is transparent and clear to the user, and explains why certain types of data are needed for the normal operation of the system [10].

In this paper, we extend the existing smart home solution with the cloud service responsible for GDPR-compliant data handling. This service allows administrators to handle privacy policy updates, and the users to request the export or deletion of personal data, as well as the deletion of the user account. First, we introduce the smart home solution architecture in Section II. Then, in Section III the operation of the GDPR service is explained, while the results of functional and performance testing are presented in Section IV and Section V.

II. Smart Home System Architecture

The smart home solution we extend is comprised of a gateway, client applications (Android, iOS and web) and cloud services.

The gateway is a key component in the smart home because it acts as a bridge between clients and smart devices in the smart home system. Gateway's main purpose is to pull together all compatible devices into a universal platform. This allows applying control scenarios to all of them while being agnostic of the actual communication interface – ZigBee, ZWave, and IP nodes are seamlessly integrated into one unified device/node network. On top of this core functionality gateway implements network API's for client applications, mechanism to define and execute rules, advanced control over the home zones, firmware upgrade, backup/restore, etc.

Sandra Bugarin is with OBLO Living, Narodnog fronta 21a, Novi Sad, Serbia (e-mail: sandra.bugarin@ oblolіving.com).

Sandra Ivanović is with the Faculty of Technical Sciences, University of Novi Sad, Serbia (e-mail: sandra.ivanovic@rt-rk.uns.ac.rs).

Marija Antić is with the Faculty of Technical Sciences, University of Novi Sad, Serbia (e-mail: marija.antic@rt-rk.uns.ac.rs).

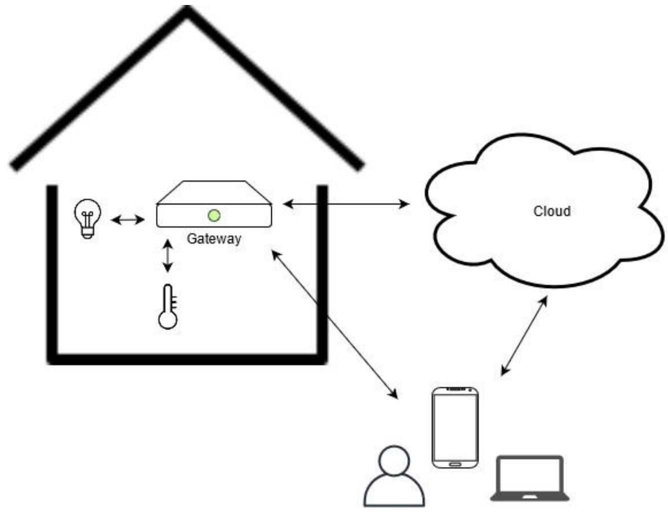


Fig. 1. Smart home system architecture.

Client applications provide the user interface for system provisioning, configuration and management. They enable users to access their home gateway in local network or remotely over the cloud. On the other hand, the cloud is responsible for user and gateway identity management, mirroring the smart home gateway configuration and data, allowing remote access for the client applications, historical data collection and analytics.

For certain functionalities of the system to be enabled, the user needs to provide the address of the household, i.e. their geolocation [11]. Also, phone number and email are needed for the purpose of smart notifications. Additionally, the system collects and stores the state changes of all devices, in order to provide the users with the possibility to inspect the way certain device types have been used in the previous period [12]. Also, the information about the local IoT network is stored for diagnostics purposes. All of these entries represent the data that should be treated according to GDPR.

III. GDPR-Compliant Data Handling

To comply with the GDPR requirements, the microservice is created within the smart home solution cloud, which enhances the system with the following functionalities:

- Update of Privacy Policy and Terms of Service
- Export of Personal Data,
- Deletion of User Account

In this section, the details about the service implementation will be presented. All of the cloud services are highly available, and GDPR service is no exception. The simplified architecture is presented in Fig. 2. Multiple instances of the service implemented in Node.JS are running on the environment. They share the long-term MongoDB data storage, as well as the temporary Redis storage. Also, the shared cloud storage disk is available to all services that need to store large files, not suitable for MongoDB database. To control the load and orchestrate tasks within the environment, RabbitMQ is used.

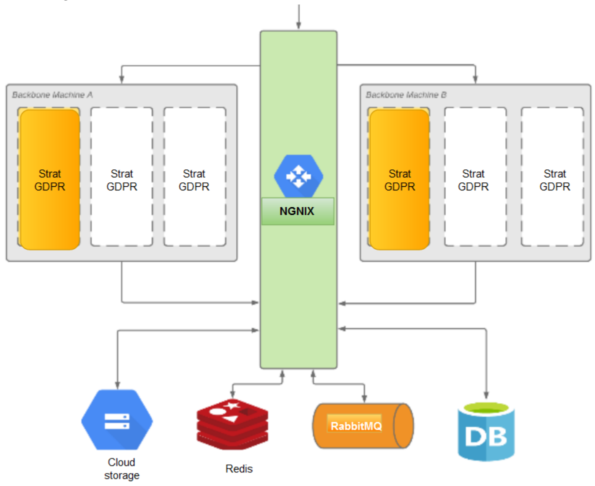


Fig.2. Highly available GDPR service.

A. Update of Privacy Policy and Terms of Service

The Privacy Policy should help users to understand what information is collected, for which purpose, and how users can update, export, and delete their information. Information about privacy policy and terms of service is the part of registration process, so all new users have to read it, and agree in order to register their smart home account.

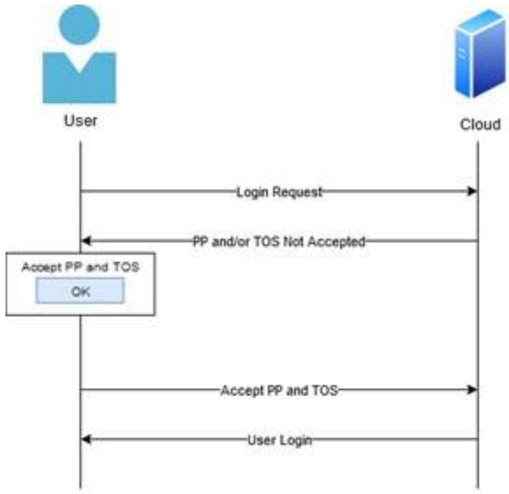


Fig. 3. Privacy policy acceptance upon login.

The existing users that have not read the new privacy policy well be prompted to read it and agree to it after logging onto web or mobile applications, as presented in Fig. 3. Until they agree to new privacy policy, users will not be able to use the applications. Gateways connected to user's account will be deactivated and prevented from sending any new sensory information to the smart home cloud.

Users can reactivate their account and gateways if they accept new privacy policy on login, or if they click on the link to new privacy policy that has been emailed to them.

B. Personal Data Export

As already said, personal data consists of user's personal profile information, such as name, email, phone contact, geolocation and address of the household. It also includes the state history of the end devices in the system, gateway backups and local IoT network history logs, which are stored for the purpose of diagnostics. Therefore, the exported data contains three groups of JSON files. The first group contains the data from the user's profile, the second one represents the snapshot of the gateway's current state, while the third one represents the usage history of all devices that have been connected to user's gateway(s).

The data export service will run on demand, under control of administrator. It performs the following tasks:

- Database crawl for personal data,
- Compression of this data to a ZIP archive, which is temporarily stored in the cloud, until the user downloads it,
- Deletion of outdated personal data

The collection and deletion of all data for an individual user can be started by the administrator, upon a request from the user (Fig. 4). Administrators can start or stop data export task that have not been completed, and delete completed export tasks and data file associated with them if they are older than 15 days. Also, they can monitor the progress of currently running data collection tasks. Administrators are not able to view the contents of the exported data files or to remove data export tasks that still haven't completed.

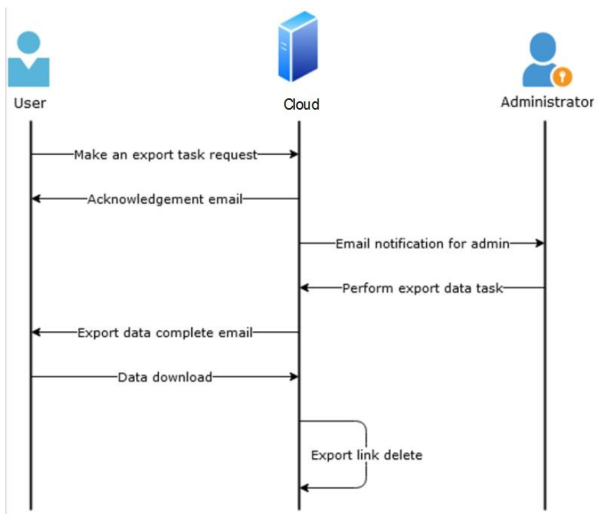


Fig. 4. Export data flow

When the user makes a personal data export request, they will be notified via email that their request has been acknowledged. A similar notification will be sent to the administrator, with the link that allow to monitor that export task. When the export task has been completed, another notification will be sent, this time with a HTML link to data export file. Download link will be available for the next 15 days. After this period, the export task and the associated export data file will be removed.

By default, only single process per backbone instance is allowed to execute data collection and compression tasks. Reason for this is intense I/O and CPU utilization (for DB crawl and data compression, respectively). Every process will be given a certain amount of time to complete it (e.g. 5 minutes) by placing the key-value pair in redis with the same expiration time. Given that database and redis are the only shared state between backbone instances, they can be used for tracking of task progression: if one of the instances that is running collection task crashes or restarts, time for task completion will expire and this task will fail.

C. User Account Deletion

At any time, a user can request to delete their account. Administrators are obliged to fulfil this request, by performing the account deletion operation via the administrative portal – Fig. 5. During this process, all of the gateways assigned to the user will be un-assigned from the user account, and all personal data from the cloud will be deleted.

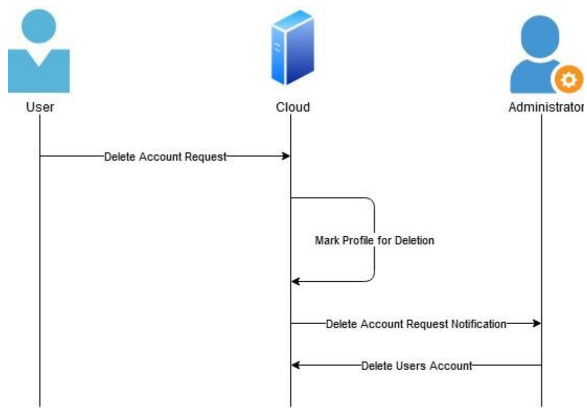


Fig. 5. Account deletion flow

However, the device usage data will be kept for analytics purposes. This data is in anonymized state, which means that it does not contain any information that can be traced to the original user.

IV. Functional Verification

A. Privacy Policy Acceptance and Modification

During the process of account creation, the user is asked to agree to the terms of service and the privacy policy.

The administrator can upload the new privacy policy and terms of service documents using the web portal for system administration – Fig. 6.

Fig. 6. Privacy policy and terms of service update.

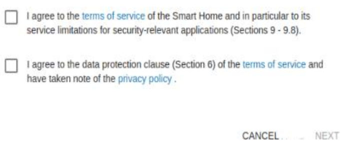


Fig. 7. Modal dialog prompting the user to accept new privacy policy.

On next login attempt, every user will be prompted to accept new privacy policy and terms of service via modal dialog – Fig. 7. Until they accept, they will not be able to use the applications.

B. Data Export

On the user profile, a button is implemented which allows them to request the export of personal data. This button is disabled if another request is already processed. This tab also contains a link to personal data when collection task is finished – Fig. 8. Implications are, that a new data export request can be made after 15 days (guaranteed duration of the valid export link) plus the time needed to perform the data export request.

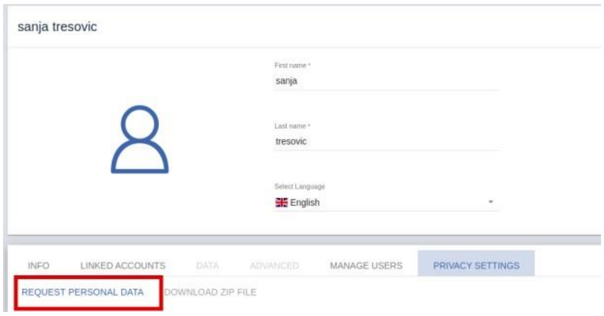


Fig. 8. User requesting personal data export.

From the administrator side, the status of the pending, current and past data export tasks can be monitored, as in Fig. 9.

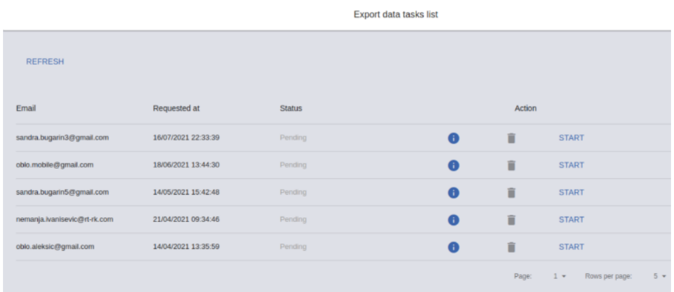


Fig. 9. Administrative panel for export task monitoring.

V. Performance testing

We have tested the implemented solution to asses the average time needed to prepare the export ZIP file with user data, depending on the data size. Typically, the size of the exported data is 5-10 MB, although for the setups with many devices it can increase up to 30 MB. The time needed for data export is presented in Fig. 10. It can be observed that the data export can be performed in less than 10 s for typical setups, while for the larger setups the time needed increases to the order of minutes. However, since the user will be informed by the notification when this process is finished, the performance of the solution is acceptable for the practical purposes.

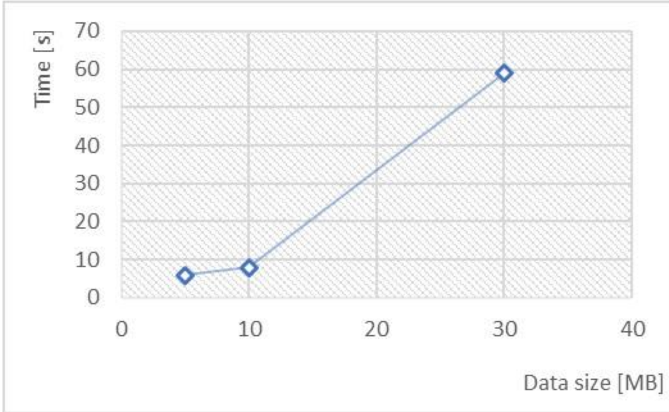


Fig. 10. Time needed to export data depending on the total size of the data file.

VI. Conclusion

In this paper, one implementation of the GDPR-compliant data handling in smart home solution has been presented. The cloud service was created, that handles the relevant aspects of data handling and user consent management, such as the update of terms of use and privacy policy, data export and account deletion. The implemented functionality has been verified, and it has been shown that the times needed for data export are acceptable. In the future work, this solution will be extended to allow users the finer granulation over the types of data collected and services enabled. For example, the users may want to opt out of the advanced functionalities, based on data analytics and machine learning, while still wishing to allow the exchange of data needed for the basic system operation.

Acknowledgment

This research has been supported by the Ministry of Education, Science and Technological Development through the project no. 451-03-68/2020-14/200156: “Innovative scientific and artistic research from the FTS activity domain”.

References

[1] A. Tsohou, E. Magkos, H. Mouratidis, G. Chrysoloras, L. Piras, M. Pavlidis, J. Debussche, M. Rotoloni, B. Gallego-Nicasio Crespo, “Privacy, security, legal and technology acceptance elicited and consolidated requirements for a GDPR compliance platform,” *Information and Computer Security*, vol. 28, no. 4, pp. 531-553, Oct. 2020

[2] P. Porambage, M. Ylianttila, C. Schmitt, P. Kumar, A. Gurtov, A. V. Vasilakos, “The Quest for Privacy in the Internet of Things,” *IEEE Cloud Computing*, vol. 3, no. 2, pp. 36-45, Mar. 2016

[3] S. Wachter, “The GDPR and the Internet of Things: a three-step transparency model,” *Law, Innovation and Technology*, vol. 10, no. 2, pp. 266-294, Sept. 2018

[4] S. Watcher, “Normative challenges of identification in the Internet of Things: Privacy, profiling, discrimination, and the GDPR,” *Computer Law & Security Review*, vol. 34, no. 3, pp. 436-449, June 2018

[5] C. Perera, R. Ranjan, L. Wang, S. Khan and A. Zomaya, “Big Data Privacy in the Internet of Things Era,” *IT Professional*, vol. 17, no. 3, pp. 32-39, May 2015

[6] A. D. Kounoudes, G. M. Kapitsaki, “A mapping of IoT user-centric privacy preserving approaches to the GDPR,” *Internet of Things,* vol. 11, no. 100179, Sept. 2020

[7] H. Lee, A. Kobsa, “Understanding user privacy in Internet of Things environments,” *Proc. of World Forum on Internet of Things (WF-IoT),* Dec. 2016

[8] Moataz Soliman, Tobi Abiodun, Tarek Hamouda, Jiehan Zhou, ChungHorng Lung, “Smart Home: Integrating Internet of Things with Web Services and Cloud Computing,” *International Conference on Cloud Computing Technology and Science,* Dec. 2013

[9] S. Zheng, N. Apthorpe, M. Chetty, N. Feamster. “User Perceptions of Smart Home IoT Privacy”, *Proc. of the ACM on Human-Computer. Interactiion.* vol. 2, no. CSCW, pp. 1-20, Nov. 2018

[10] K. Renaud,L. A. Shepherd, “How to make privacy policies both GDPRcompliant and usable,” *International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA)*, Nov. 2018

[11] M. Matić, M. Tucić, M. Antić, R. Pavlović, “Using online third party geolocation services to improve smart home user experience,” Serbian Journal of Electrical Engineering vol. 17, no. 1, pp. 83-94, Feb. 2020

[12] S. Ivanović, M. Antić, I. Papp, N. Jović, “Data Acquisition, Collection and Storage in Smart Home Solutions,” *Proc. of 6th International Conference on Electrical, Electronic and Computing Engineering (IcETRAN)*, May 2019

Combined adaptive load balancing algorithm for parallel applications

Luka Filipović, Božo Krstajić, *Member IEEE*, Tomo Popović, *Senior Member IEEE*

***Abstract*— Development and improvement of efficient techniques for parallel task scheduling on multiple cores processors is one of the key issues encountered in parallel and distributed computer systems. The purpose of process distribution improvement in parallel applications is in increased system performance, reduced application execution time, reduced losses and increased resource utilization.**

This paper presents combined adaptive load balancing algorithm based on domain decomposition and master-slave algorithms and its core scheduling adaptive mechanism that handles load redistribution according obtained and analyzed data. Selection of distribution algorithm, based on collected parameters and previously defined conditions, proved to deliver increased performances and reduced imbalance. Results of simulations confirm better performance of proposed algorithms compared to the standard algorithms reviewed in this paper.

***Index Terms*— parallel programming, load balancing algorithm, tash scheduling, adaptive algorithm.**

I. Introduction

Distributed computer systems enables the delivery of computing resources necessary to solve complex problems with requirements that exceed the capabilities of the most powerful personal computers. High-performance computers, as one of the powerful elements of distributed computer systems, lead to complex solutions by using computer simulations enabling progress in all scientific fields. Parallel processing supports execution of several processes and instructions simultaneously, with a goal to save time and execute faster and more efficient complex applications in scientific and industrial applications. [1] [2].

The focus of many researches in the parallel processing field is process of finding optimal distribution of tasks in order to increase efficiency, reduce execution time of parallel applications and reduce communication time of computer resources. In order to achieve the highest parallel application efficiency, it is crucial to optimize the assignment of tasks to parts of the distributed computer system (cluster nodes and its CPU cores) and monitor their execution.

The subject of this research was combined adaptive algorithm (CAA) [3][4], which uses combination the static and dynamic load balancing algorithms to improve the performance of independent parallel tasks scheduling without significantly complicating the whole process. It uses an adaptive innovative mechanism for choosing load balancing algorithm for distribution of unexecuted autonomous tasks depending on the segments in which losses are the least and by limiting the algorithm at times when it causes losses.

II. Load balancing algorithms

Load balancing in parallel processing is defined as process of achieving parallelism by redistributing the load of parallel segments during the execution of a parallel program [5] [6]. The primary goal of load balancing algorithms is to find the optimal execution schedule that defines the initial execution time and the execution order of all tasks that run on a particular resource. Load balancing of parallel applications is process of reducing computation time achieved by reducing communication time, synchronization time between processes and waiting time due to uneven process distribution [7].

The imbalance of parallel applications most often occurs due to uneven load between cores, excessive communication between cores or waiting of group of cores for others to finish assigned jobs [8]. In a real distributed environment, resource load varies over time and it is not always possible to improve the use of resources that are completely free or equally loaded. It is not possible to determine or predict the length of processes that run on separate computers or delays due to communication between computers. Therefore, there is a longer execution of the parallel application and a decrease in resource utilization. The end of the execution of a parallel application or the beginning of the postprocessing phase directly depends on the execution time of the part of the application on the core that is assigned the most process or the processor with the lowest frequency.

Load balancing algorithms are divided as static and dynamic, depending on the type of job scheduling. Static load balancing algorithms have good usability and efficiency on homogeneous clusters while they execute tasks on all cores which have similar duration. Performance of programs using these algorithms is reduced at the end of the runtime without possibility of rescheduling. One of widely used static algorithms is domain decomposition algorithm. On the other side, dynamic algorithms can give better efficiency on heterogeneous system, but make unnecessary communication during executing time. The master slave algorithm is a one of the typical representatives of dynamic algorithms. Domain decomposition and master-slave algorithms have their advantages and disadvantages depending on the characteristics of the resource, the specific parallel application for which load balancing is performed and the duration of processes that are executed in parallel [9-11].

Luka Filipović is with the University Donja Gorica, Oktoih 2, Podgorica, Montenegro (e-mail: luka.filipovic@udg.edu.me).

Božo Krstajić is with the Faculty of electrical engineering, University of Montenegro, Džordža Vašingtona bb, Podgorica, Montenegro (e-mail: bozok@ucg.ac.me).

Tomo Popović is with the Faculty of Information Systems and Technologies, University Donja Gorica, Oktoih 2, Podgorica, Montenegro (e-mail: tomo.popovic@udg.edu.me).

Adaptive algorithms are advanced dynamic algorithms with adaptive strategy for task distribution scheme that is activated depending on the load change of the distributed system during operation.

III. Combined Adaptive Load Balancing Algorithm

The combined adaptive algorithm (CAA) is successor an improved version of combined algorithm (CA) [12]. It presents an adaptive decision model that selects an adequate algorithm based on data on the state of the resource on which the parallel application is running and the duration of finished tasks.

In the preprocessing phase, as in the CA algorithm, the input data is divided and tasks are prepared for execution. Before starting parallel simulations, the analysis of the distributed resource configuration is performed and the obtained data are used in the later analysis.

In the parallel processing of the combined adaptive load balancing algorithm, three execution phases stand out (Figure 1):

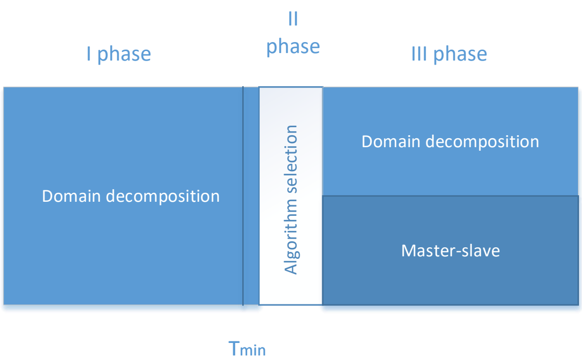


Fig 1. Execution phases of the proposed CAA algorithm

In the first phase of the combined adaptive algorithm, the domain decomposition algorithm is executed. It has the highest efficiency and the lowest losses in the initial phase of program execution. The algorithm stops working when the first (“fastest”) core completes the assigned job (T_{min}) and sends instructions to the other cores to stop working after completing the task they are processing at that point. The described procedure reduces the losses of the first execution phase to a minimum.

In the second phase of the algorithm, based on the amount and duration of performed tasks, cluster configuration and its load, an adaptive approach is used to select the algorithm for the scheduling of the remaining tasks in third phase. Upon initiating an interrupt at the end of the first phase, each CPU core sends to a predefined core a data containing the duration of the performed tasks. The predefined core receives the sent data and processes them, making an array with the number of executed tasks for each core and through executed and unexecuted tasks and selects the algorithm to be executed in the third phase according to the defined decision algorithm. The decision on the algorithm in third phase is made on the basis of the following parameters:

- the homogeneity of allocated resources,
- the total number of assigned cores,
- the numbers of completed tasks for each core individually and
- the execution time of each task individually.

The homogeneity of the allocated resources (the examination of whether they can be considered homogeneous or heterogeneous) is performed by comparing the performance values of the allocated nodes of the distributed resources. A measure of the performance of an individual resource can be core frequency, node memory or node network speed. Depending on the architecture of the distributed system and the type of tasks, one or more node performance measures can be taken. In the presented research, the core frequency (Hz) was used as a measure of the node performance of the distributed system.

The total number of assigned cores is defined when the application is started.

The number of completed tasks per core represents the part of the total number of tasks performed up to the moment T_{min}, when the first core performed the assigned tasks and initiated the interrupt, for each core separately. The data is expressed as a sequence whose number of elements is equal to the number of assigned cores, and the elements are the numbers of completed tasks for each core individually. The total number and type of tasks depends on the parallel application being executed and the input data, and the division is done before the parallel processing.

The execution time of each individual task is a matrix that contains the data on which core the task was executed and the duration of each task (ms) that was completed.

Based on the above parameters, the conditions for selection of an adequate distribution algorithm in the third phase can be defined. These conditions are defined by variables U_i that have binary values. Thus, the variable U_i takes the value 1 if the i-th condition is met, and otherwise U_i takes the value 0.

The first and eliminatory condition (U_e) for the selection of the distribution algorithm is the condition that the remaining number of tasks is less than or equal to the number of available cores. If the conditions U_e ($U_e = 1$) are met, the DD algorithm is selected for execution in the third phase, ie each of the remaining tasks is assigned one core for execution.

If the eliminator condition is not met ($U_e = 0$), the choice of algorithm is made based on a combination of the following conditions:

- U_1 - cluster homogeneity condition: this condition is fulfilled (U_1=1) if CPU cores of the same or approximate operating clock are assigned, ie. if the standard deviation of the operating clock of all cores is less than the set value;
- U_2 - number of cores condition: this condition is fulfilled ($U_2 = 1$) if the number of cores is less than a predefined number of cores, ie if the losses of the master core in the MS algorithm cannot be ignored;
- U_3 - condition of uniformity of the number of performed tasks: this condition is fulfilled ($U_3 = 1$) if the number of performed tasks for each core is approximate, ie. if the value of the standard deviation of the number of completed tasks per core is less than the predetermined value;
- U_4 - condition of uniformity of duration of performed tasks: this condition is fulfilled ($U_4 = 1$) if the duration of performed tasks per core is approximate, ie. the

value of the standard deviation of the execution time of each task per core is less than the predefined value.

The decision algorithm checks the fulfillment of conditions that depend on the values of the parameters. Choice of the algorithm itself adapts to the current performance of the allocated resources and the state of the performed tasks in the first phase. Thus, the proposed adaptive algorithm determines whether the domain decomposition or master-slave algorithm will be executed in the next phase based on the fulfillment of the defined conditions according to the principle: the more conditions are met, it determines the choice of DD algorithm in the third phase and vice versa.

In order to enable additional adaptation of the decision algorithm to a specific application and distributed system, each of the conditions can be weighted with real coefficients K_i, $K_i \in [0,1]$ which enables the exclusion of some conditions or assigning greater or lesser importance to some of the conditions. This does not apply to an eliminatory condition that is considered independently of the other conditions. The coefficients K_i are assigned a maximum value of 1 if this condition is fully taken into account, while $K_i = 0$ excludes the influence of this condition from the influence on the choice of algorithm. Coefficients should be defined separately for each application and distributed resource depending on previously obtained results and experiences.

Finally, based on the above conditions, we can define the decision function on the basis of which we select the algorithm in the third phase:

$$U = \sum_{i=1}^{4} Ki * Ui. \quad (1)$$

The threshold value of the decision function U should also be defined, on the basis of which one or another algorithm is selected for the third phase (DD or MS). Since the maximum of the function U is achieved by the fulfillment of the conditions K_i*U_i and that determines the choice of the DD algorithm, then half of the maximum value of the function U is taken as the threshold value, ie

$$\mathrm{P} = \frac{\sum_{i=1}^{4} \mathrm{Ki}}{2}. \quad (2)$$

Therefore, if it's satisfied

$$U \geq P \quad (3)$$

it is necessary to select the DD algorithm in the third phase or the MS algorithm if condition is not satisfied.

Figure 2. shows a schema of the decision making process for the selection of algorithm in second phase. As already mentioned, based on the presented parameters, defined conditions and coefficients, the algorithm for the distribution of tasks in the third phase is selected.

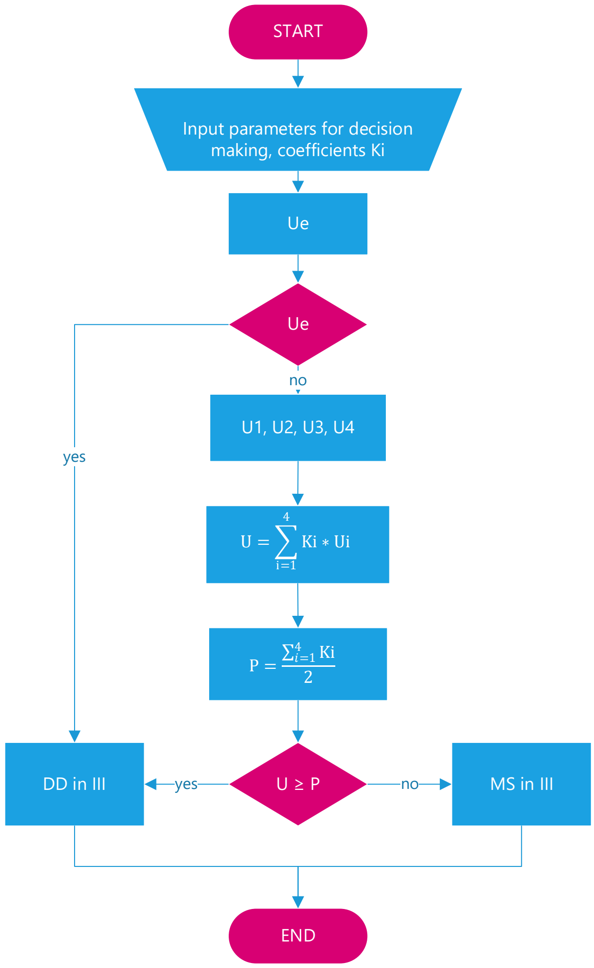


Fig 2. Scheme of the decision making process for the selection of algorithm in Phase II

The selected algorithm (DD or MS) is executed in the third phase.

If the DD algorithm is selected, each core receives a portion of the list of unfinished tasks. Each core gets assigned one of the remaining tasks to solve if the remaining number of tasks is less than or equal to the available number of cores (condition Ue). Otherwise, the number of assigned tasks for each core is determined in proportion to the number of tasks completed in the first phase on each core separately.

In the case of selecting the MS algorithm, the core that performed the analysis in the second phase is determined as the master core. It contains information with a list of all unfinished tasks that are assigned to slave cores for execution in the third phase of the algorithm.

The proposed CAA algorithm will increase efficiency and shorten the execution time of parts of a parallel application in the third phase according to the interruption of the execution of the first phase, the analysis of the state of resources, the adaptation from the second phase and the redistribution of tasks.

The efficiency of the CAA algorithm has been improved due to process reallocation, reduced kernel latency for new instructions, and improved resource utilization by adapting the allocation to the distributed system architecture and application-specific. Therefore, the execution time of the proposed algorithm will be shorter than the execution time of the standard DD algorithm if measured under the same conditions. The CAA algorithm is similar to the CA algorithm in the case of deciding that a dynamic process allocation along with the MS algorithm is required in the third stage.

The disadvantages of the proposed CAA algorithm are the interruption of task execution at the end of the first phase and the duration of adaptation in the second phase. Interrupting the execution of tasks in the first phase may increase the duration of this phase if there are one or more tasks whose duration is significantly longer than the duration of other tasks. This phenomenon would cause an increase in the duration of the first phase, which may affect the performance of the entire algorithm. In that case, the efficiency would be the same as with the classical DD algorithm. The second phase, due to its short duration, cannot significantly affect the overall efficiency of the parallel application.

The proposed CAA works as a DD algorithm during the period of its maximum efficiency and stops working when its efficiency starts to decline. The proposed adaptive algorithm will have a significantly better performance than the domain decomposition algorithm in the case when the basic algorithm has low efficiency due to interruptions and redistribution of tasks.

The CAA algorithm will have better performance than the MS algorithm because the MS algorithm does not execute tasks on the master core and generates more communication losses than the proposed CAA algorithm. The MS algorithm will have lower efficiency than the proposed algorithm because it starts as a DD algorithm and redistributes and selects the algorithm for execution based on parameters in order to achieve better use of resources and efficiency.

In case of large losses during third phase, it is possible to re-initiate the interruption and repetition of the decision algorithm, ie adaptation based on new parameters, re-selection of the algorithm and its start to get the best use of resources.

IV. THE ANALYSIS OF SIMULATION RESULTS

For the purposes of research and testing of the subject algorithms, a parallel version of the crossbar commutator performance simulator (CQ) [13] was used, as a numerically demanding example of a parallel application with several independent processes. The algorithms were tested on different distributed computing environments and run under different resource loads. Each simulation was performed ten or more times and the averaged results of the execution time are presented here. The performance of the combined adaptive algorithm was verified on the example of a 16-port CQ simulator with 1,000,000 requests and 3072 generated tasks. Simulations performed on the Paradox HPC cluster of the Institute of Physics in Belgrade. At the time of the simulation, the cluster consisted of 106 computing nodes based on two octa-core Xeon 2.6GHz processors with 32GB of RAM and NVIDIA® Tesla ™ M2090 cards. The performance of the combined adaptive algorithm is compared with the performance of the algorithms that make it up. Simulations were performed on 16, 32, 64 and 128 cores. The input files were copied to the nodes on which the simulations were run in the preprocessing phase, thus reducing the impact of communication between the nodes.

In the presented simulations, the value of standard deviation 10% of the average value of the core operating clock was used for condition U_1. A threshold of 32 cores is defined for condition U_2. For conditions U_3 and U_4, the value of the standard deviation is 25%. The coefficients used in these simulations are $K_1 = 0$, $K_2 = 1$, $K_3 = 1$ and $K_4 = 0.5$. Priority in decision making is given to the number of cores on which the simulation is performed and the number of performed tasks per core. A lower priority was given to the duration of the tasks, and due to the coefficient $K_1 = 0$, the influence of cluster homogeneity was not taken into account. The average results of parallel application execution with DD, MS and CAA algorithm for different number of used cores are shown in Figure 3.

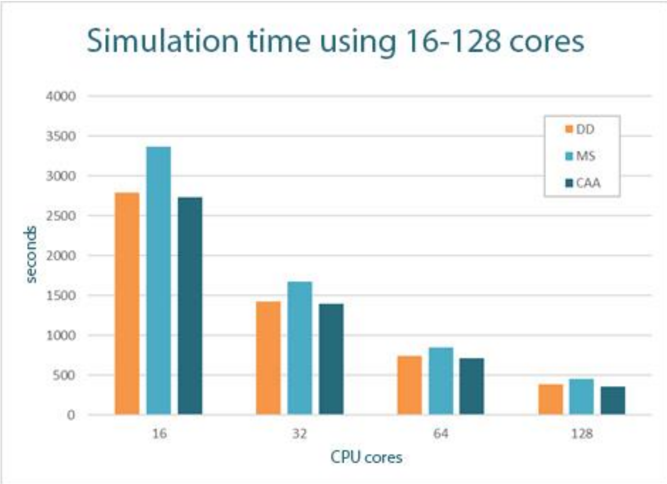


Figure 3. Average execution time of simulations using DD, MS and CAA algorithms on 16-128 cores

The combined adaptive algorithm completed simulations faster than the domain decomposition and master-slave algorithms in all conditions. The best results and the greatest benefits due to the redistribution of tasks were determined in cases of performing simulations on a number of cores. The simulations showed the longest execution time with the master-slave algorithm, especially on a small number of cores due to its previously described shortcomings.

The domain decomposition algorithm performed simulations faster than the master-slave algorithm. The input data was transferred before the simulations and most tasks were performed at approximately the same time, as shown in Figure 3. Therefore, the static distribution proved to be sufficient and the domain decomposition algorithm showed better performance than the master-slave algorithm.

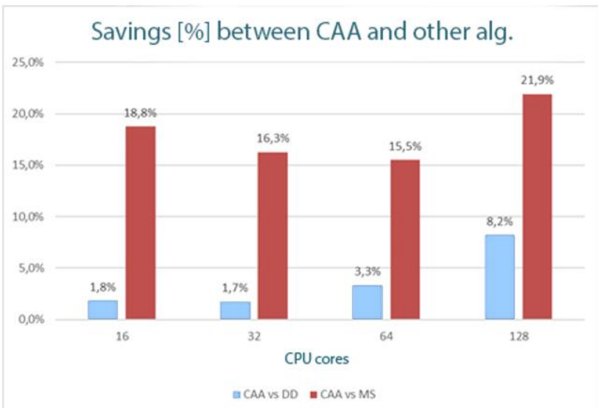


Figure 4. Savings during algorithm execution and comparison between combined algorithm and domain decomposition and master slave

Figure 4 shows the execution time savings between the combined adaptive algorithm and the algorithms that make it up. The domain decomposition algorithm required more time than the combined adaptive algorithm due to the static distribution throughout the execution process. The difference between the combined adaptive and domain decomposition algorithms ranges from 1.7% to 8.2%. The biggest difference was recorded when executing the application on 128 cores.

The differences between the combined adaptive algorithm and the master-slave algorithms are due to the loss of the master-slave algorithm due to the distribution of tasks and communication between cores during the entire program execution process. The execution time difference between the combined adaptive and master-slave algorithms ranges from 15.5% to 21.9%. The inability to execute tasks on the master core produced losses during execution on a smaller number of cores. Increased communication between cores throughout the execution of the simulation caused the largest difference between the results listed on 128 cores.

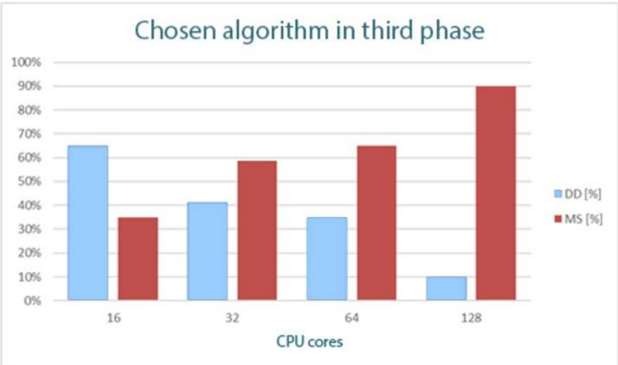


Figure 5. Selected algorithm in the third phase of CAA

Figure 5. shows the results of the selection of the algorithm in the second phase according to the received and analyzed data and the decisions made at the end of the second phase. The domain decomposition algorithm was chosen in most cases when the simulation was performed on 16 cores, because the execution was detected on less than 32 cores and an even number of tasks that needed to be redistributed. On the other hand, master-slave was chosen in cases of simulations on 32 or more cores because the decision algorithm from the second phase based on parameters discovered the number of available cores, different number and duration of performed tasks and selected this dynamic algorithm for the third phase.

V. Conclusion

The paper presents an original adaptive load balancing algorithm for parallel applications that combines the operation of static and dynamic algorithms. Domain decomposition and master slave algorithms were used on the basis for the proposed algorithm, as one of the most common algorithms in practice. As none of the algorithms provides good results in a wide range of applications and types of distributed systems, the following research was based on the idea of combining the mentioned algorithms in order to improve the parallelization performance without complication of the algorithm. Based on the identified advantages and disadvantages of standard algorithms, a combined adaptive algorithm is proposed. The idea of combined algorithms is to work in the phases when composite algorithms have the best performance. The advantages of the proposed solution are following:

- improved parallel application efficiency and cluster utilization in relation to basic algorithms due to task redistribution and reduced execution time;
- parameters and conditions for the selection of algorithms have been identified according to the status of resources and the point of execution of the application and determine a more adequate static or dynamic distribution of the process by an adaptive strategy
- weighting coefficients (K_i) adjust the adaptive load balancing algorithm and parallel application to the infrastructure
- applicability of the proposed adaptive part of the decision algorithm is possible in any load balancing algorithm and
- the proposed algorithm is applicable to all parallel applications consisting of several independent tasks.

The paper presents the results of executing domain decomposition, master-slave, combined and combined adaptive algorithm on different computer resources with the help of numerically demanding parallel application of CQ simulator. Comparison of the results of simulations with different loads and configurations of distributed resources confirms the better performance of the proposed algorithm in relation to the basic algorithms considered in the paper.

Acknowledgment

This research is supported in part by the EuroCC project, grant agreement grant agreement 951732 EuroCC-H2020-JTI-EuroHPC-2019-2.

References

[1] S. Tanenbaum and M. van Steen, Distributed Systems: Principles and Paradigms, 2nd Edition, Pearson Education. Inc., 2007.
[2] B. Barney, Introduction to Parallel Computing, Lawrence Livermore National, 2012.
[3] L. Filipovic, “Combined adaptive load balancing algorithm for parallelization of applications”, PhD thesis, University of Montenegro, Faculty of Electrical engineering, 2019.
[4] L. Filipovic and B. Krstajic, "Combined load balancing algorithm in distributed computing environment," Information Technology and Control, vol. 45, no. 3, pp. 261-266, 2016.
[5] H. D. Karatza and R. C. Hilzer, "Parallel Job Scheduling in Homogeneous Distributed Systems," Simulation, vol. 79, 2003.
[6] J. Dinan, S. Olivier, G. Sabin, J. Prins, P. Sadayappan and C.-W. Tseng, "Dynamic Load Balancing of Unbalanced Computations Using Message Passing," in Parallel and Distributed Processing Symposium, 2007, IPDPS 2007, IEEE International, Long Beach, CA, USA, 2007.
[7] T. Rauber and G. Rünger, Parallel Programming for Multicore and Cluster Systems, Springer, 2010.
[8] D. Thiébaut, Parallel Programming in C for the Transputer, 1995.
[9] V. Sarkar. Partitioning and Scheduling Parallel Programs for Multiprocessors. MIT Press, 1989.
[10] W. D. Gropp. Parallel Computing and Domain Decom-position. In: Fifth Conference on Domain Decompo-sition Methods for Partial Differential Equations, 1990, pp. 249-361.
[11] S. Sahni. Scheduling Master-Slave Multiprocessor Systems. IEEE Transactions on Computers, 1996, Vol. 45, No. 10, 1195-1199.
[12] L. Filipovic, B. Krstajic. Modified master-slave algorithm for load balancing in parallel applications. ETF Journal of Electrical Engineering, 2014, Vol. 20, No. 1, 74-83.
[13] M. Radonjic and I. Radusinovic, "CQ Switch Performance Analysis from the Point of Buffer Size and Scheduling Algorithms," in Proc. of 20th Telecommunication Forum TELFOR 2012, 2012

A Tool for Sentence Syntax Structure Markup for The Serbian Language

Teodora Đorđević, Suzana Stojković *University of Niš, Faculty of Electronic Engineering*

Abstract— **Syntax analysis is an important part of natural language processing. The biggest challenge to defining a natural language syntax analyzer is the inability to define unambiguous formal grammars that describe the language. Because of this, rule-based syntax analyzers need to be enhanced using statistics to allow us to predict which syntax tree is most likely. In order to do this, a corpus of tagged sentences in the target language is needed. The creation of this corpus is long and tedious work. Because of this, this paper implements a visual tool for creating such a corpus for the Serbian language. A component of this tool is the syntax analyzer, which generates all the possible syntax trees based on the defined grammar such that an expert may choose one of them. The expert may also create entirely new syntax trees.**

Index Terms—**Natural language processing (NLP); Syntax analysis; CYK; Annotated syntactic corpora; Serbian language**

I. Introduction

Natural language processing is a branch of computer science that teaches computers to understand and manipulate human language. Natural language processing is a combination of computer science, linguistics and machine learning. Many NLP techniques are already developed and applied for the English language but applying those techniques to different languages can be quite a challenge. Serbian language is under-researched in the context of natural language processing. Since the Serbian language and the English language do not belong to the same language group, many approaches designed for the English language need to be significantly modified in order to be used in the Serbian language or cannot be used at all.

Syntax analysis or parsing, in general, is the process of analyzing character strings according to the rules of a given formal grammar. It is typically encountered in fields of natural languages, computer languages or data structures. In Natural language processing, syntax analysis is one of the most important phases because it builds a great foundation to natural language understanding. Syntax analysis decides whether a sentence written in natural language conforms to the rules of a formal grammar and thus whether a sentence is valid or not. Designing a quality syntax parser is extremely significant for designing a semantical analyzer, since syntax parsing precedes semantic analysis. Also, syntax analysis has its own role in Rule-based Machine Translation, Information Extraction, Question Answering systems, etc.

Teodora Đorđević is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia (e-mail: teodora.djordjevic@elfak.ni.ac.rs).

Suzana Stojković is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia (e-mail: suzana.stojkovic@elfak.ni.ac.rs).

This paper describes designing a graphic tool for syntax analysis of the Serbian language based on the syntax analyzer designed in [1]. This tool is implemented as a web application that can analyze and visualize sentences using the implemented parser. It also enables the user to draw completely new syntax trees and save them.

II. Related Work

In linguistics, a corpus usually represents a collection of texts. The main purpose of a corpus is to be used as a tool in language study. In order to study and analyze language data, having a corpus is essential.

The English language is most widely researched language and thus the largest number of corpora exists for the English language.

Corpora contain texts that are sourced from natural contexts in order to be as close to the natural language as possible. There are many types of corpora that are used in natural language processing such as reference corpora, which are fairly balanced sets of texts that accurately describe a standard language, or specialized corpora which contain texts from a particular area, such as movie reviews, magazine texts, etc. A very important category of corpora are annotated corpora which contain additional information such as part of speech tags, lemmas, metadata, additional tags, etc. Annotated corpora can thus be used in supervised learning scenarios when attempting to infer this additional data based on the text given.

There are many publicly available corpora online for various languages. These corpora can be accessed either directly online via web browser, through specialized APIs to search the corpora, or can also be downloaded in their entirety.

Most modern language processing is done using computers. This means that modern corpora must be electronically readable documents. The first such document for the English language was The Brown Corpus of Standard American English [2]. This corpus consists of one million words of American English texts printed in 1961. In order to ensure high quality and to make the corpus useful for a wide range of applications, the corpus compiled texts from 15 different categories. Keeping in mind the huge increase in processing power, as well as that the Internet generates more linguistical data than ever before, this corpus is now considered small.

An example of a modern corpus of the English language that is quite big is the Corpus of Contemporary American English (COCA) [3]. COCA is probably the most widely used corpus of English, with over one billion words. Many corpora for the English language can be found at [4].

Besides the English language many other languages of the world are being researched in the field of syntax and semantic

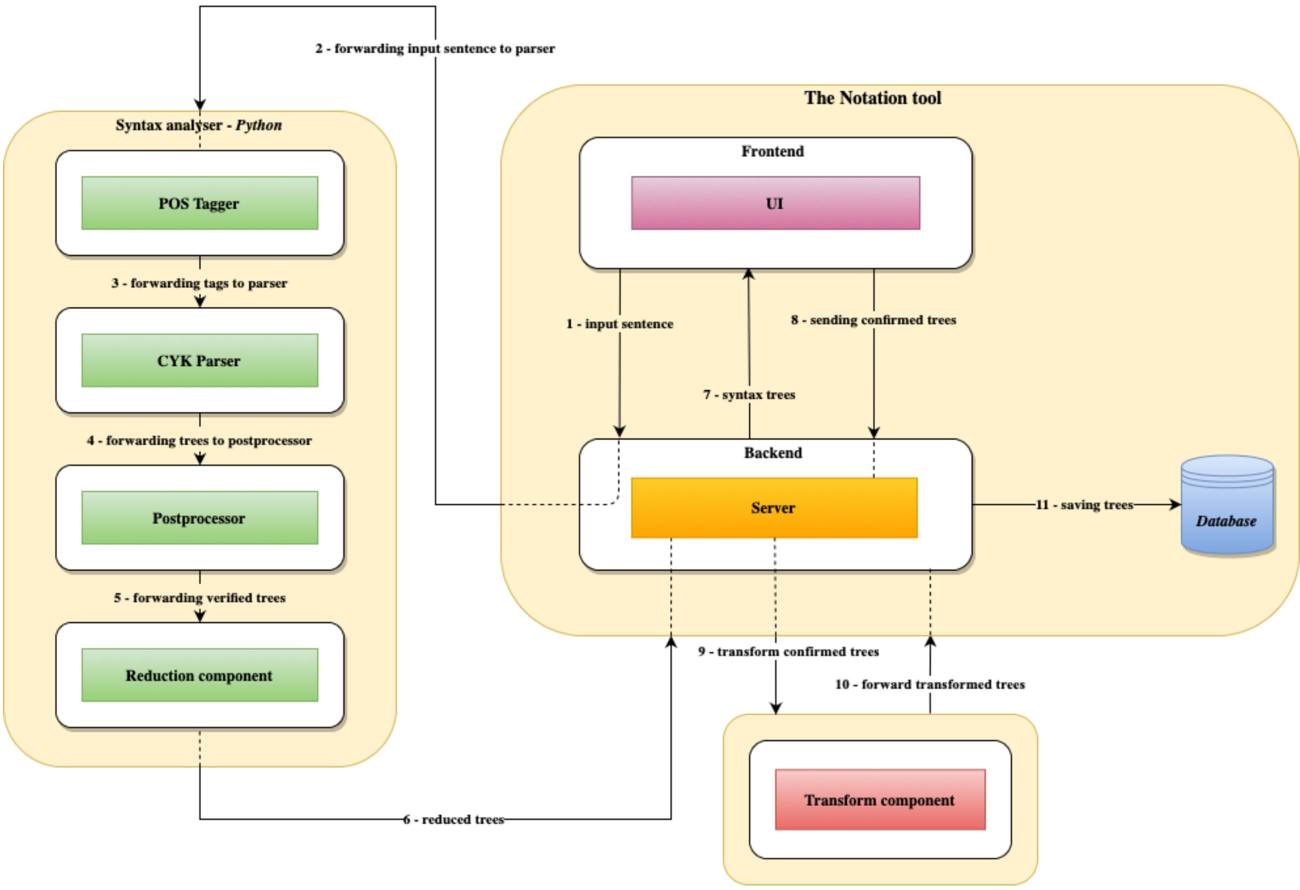


Fig. 1. The Architecture of Notation tool

analysis and are thus being compiled into language specific corpora. One such corpus is the Quranic Arabic Corpus [5], which is an annotated linguistic document, that shows the Arabic grammar, syntax and morphology for each word in the Quran. The research paper [6] describes a specialized annotated corpus for the Chinese language, used for analyzing clinical texts. This corpus is annotated with part-of-speech tags, syntactic tags, entities, assertions, and relations.

For the Serbian language, given that it is spoken by only 12 million people in the world as a first language, there is not a large number of corpora such as English or Chinese. In the last few years, there has been development of open, freely available resources and technologies for computer processing of texts in the Serbian language. This includes annotated language corpora and some of the corpora are listed below.

1. *SETimes.SR* [7] – it is based on the *SETimes* parallel corpus of newspaper articles. This is a manually annotated corpus of texts written in the standard Serbian language. This corpus is used for training and evaluation of computer models on a number of natural language processing problems. It contains 3891 sentences. The *SETimes.SR* corpus is annotated using morphosyntactic notation, lemmas, syntactic dependencies, and named entities.
2. *srWac* [8] - the Serbian web corpus, which was built by crawling the *.rs* top-level domain in 2014. It contains 555 million tokens and over 25 million sentences arranged in about 1.3 million documents.
3. *MULTEXT-East* [9] - is a multilingual dataset for language research. This project consists of mainly Central and Eastern European languages, including Serbian.
4. *ReLDI-NormTagNER-sr 2.1* [10] - is a manually annotated corpus of Serbian tweets. It is meant for use in the fields of tokenization, sentence segmentation, word normalization, morphosyntactic tagging, lemmatization and named entity recognition of non-standard Serbian.

As mentioned earlier, there is a corpus that has marked dependency syntax at the sentence level, but there is still no corpus for the Serbian language that contains fully marked syntax trees. Precisely for this reason, the idea arose to create such a tool that will enable the creation of a corpus of syntactic trees for the Serbian language.

There are some visualization tools for drawing syntax trees, but they are mostly limited to inputting a syntax tree, and then getting a visualization of that tree. The tools which expect user to enter syntax trees and then visualize it are shown here [11, 12]. The tool where the user can draw syntax tree from scratch is TreeForm [13]. This tool offers wide palette of elements that can be drawn in order to create a syntax tree. There are several simpler solutions than TreeForm, such as [14][15]. All these tools are only intended for visualizing syntax trees. They do not support using a syntax parser in the background, which would generate syntax trees based on the entered sentence as suggested in this paper.

III. The Functionality Of The Notation Tool

A. *Syntax Analyzer*

The parser that was created in [1] achieved excellent performance and performed real-time parsing. This parser consists of three components:

1. **POS Tagger** – when a sentence is forwarded to the syntax analyzer it is necessary to extract POS tags first, because a syntax analyzer can only recognize tags, not actual words. This tagger is explained in detail in [1]. The tagger returns tags that have special meaning. For example, some of the valid tags are 'nn' (noun in nominative), 'vm' (main verb), 'sl' (preposition in front of locative). Every POS of the Serbian language has its own abbreviation, where every letter has its own meaning. These tags are then forwarded to syntax analyzer, and later displayed in syntax trees above actual words of the sentence.
2. **CYK Parser** – this parser is implemented to achieve optimal performances while analyzing sentences. Also, this parser, as defined in [1], is capable of recognizing all the syntax trees, like the parser in NLTK [16], but with significantly reduced parsing time.
3. **Postprocessor** – this layer is added because the number of syntax trees that are generated based on grammar designed for CYK Parser was large. To reduce this number, a series of rules is defined. These rules eliminate syntax trees that aren't consistent with Serbian grammar. The postprocessing phase reduced the number of syntax trees by 54%.

The problem with this syntax analyzer, despite adding a postprocessing phase, is that it generated multiple trees for a single sentence. In order to solve this problem, it is necessary to add statistics that will enable to generate only one syntax tree as a result of a syntax analysis. To be able to implement statistical parsing, it is necessary to have a corpus of marked sentences, which is not the case for the Serbian language. For this reason, the idea of creating a visual tool arose. This tool will enable simple drawing and visualization of syntax trees and thus lead to the generation of a corpus of marked sentences that will be used further.

The notation tool works as follows:

1. The user enters the sentence they want to tag
2. The sentence is forwarded for processing to a parser that returns the resulting syntax trees
3. The syntax trees are displayed to the user
4. The user can choose one of the following options:
   - Select the correct tree,
   - Change the tree that is the most similar to the correct tree - by adding nodes, changing the node name, deleting nodes, and switching places with nodes, or
   - Create a new tree in case all the suggested trees are wrong
5. The correct tree is uploaded and stored in the database.

The architecture of the implemented system is shown in Figure 1.

The new component of the syntax analyzer is called reduction component. This component is added specifically for this tool.

The grammar created for the Serbian language contains a huge number of rules because the Serbian language is very complex. Considering that due to the implementation of the CYK algorithm, it was also necessary to transform the grammar so that it would be in Chomsky's normal form, a large number of auxiliary shifts were introduced. The syntax tree created in this way was too large to be displayed to the user of any system and this is the reason for introducing a reduction component.

This component aims to transform the syntax tree so that it no longer contains auxiliary rules, as well as that it does not contain shifts that have been introduced to make syntax analysis simpler and more robust.

The goal of reduction is to transform the syntax tree, generated by using a more complex grammar, into a simpler tree, corresponding to a simpler grammar. The main purpose for introducing the reduction component is to visualize the trees in a way that domain experts would expect by abstracting away implementation details. Also, reduced syntax trees are smaller and easier to display. After confirming the final tree for input sentence, it is necessary to return syntax tree to original form. This is achieved by using transform component. This component accepts syntax tree in simpler grammar and transforms that tree to original grammar. The transformed tree is then forwarded to backend application and saved in a database.

Figures 2 and 3 show how a part of the syntax tree looks with and without reduction. The reduced syntax tree is significantly smaller and thus much easier to display. An entire syntax tree without reduction would be impossible to fit in the page of the notation tool. The syntagm shown in figures 2 and 3 is "Moja divna drugarica", meaning "my wonderful friend" in Serbian.

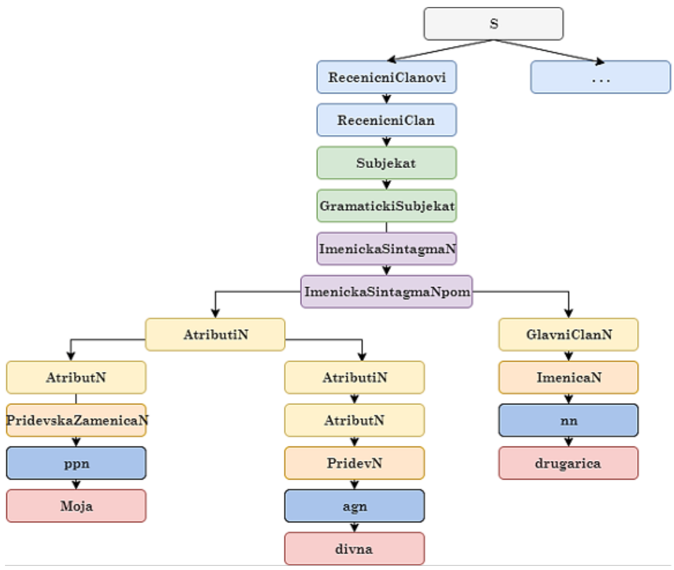


Fig. 2. Part of the syntax tree without reduction

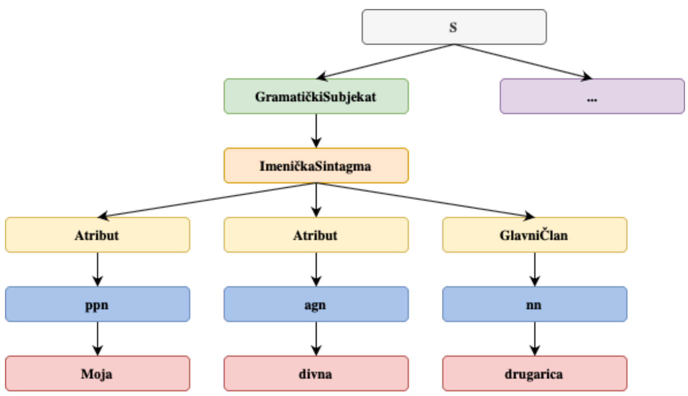


Fig. 3. Part of the syntax tree with reduction

B. *The Notation Tool*

This component is implemented as a web application so that users can use it as easily as possible. This approach was chosen to avoid any installation. The application itself is divided into three parts:

1. Frontend
2. Backend
3. Database.

The role of the Frontend is to enable:

1. **Entering a sentence whose analysis should be performed**
2. **Displaying of all syntax trees generated by the parser**
3. **Selecting a syntax tree that is correct** - the user can view a list of all syntax trees that the parser returned and check the one that is correct. This sends a request to the server with the intention to save that tree in the database.
4. **Syntax tree modification** - if the syntax tree is not completely correct, but with a few minor changes it could become correct, the tool offers the possibility to make the following syntax tree changes:
   - *adding a new node* - it is necessary to select the node to which we want to add a new descendant and select the name that will be in that node. After interacting with the component, the tree structure is automatically updated to display the changes.
   - *deleting a node* - if it is necessary to remove a node, the tool offers the option to select that node and then delete it.
   - *renaming a node* - if the tree structure is adequate and an element is incorrectly recognized and it has the wrong name, the tool allows the user to rename that node.
   - *swapping nodes* - this option exists in case the nodes are correctly recognized but they have been misplaced. It is possible to swap the places of these nodes, but only if they have the same parent. This option was introduced because a new node is always added to the end of the list of children, and if the node is deleted, a new node should be added in its place. Since the new node is always added as the last child, this functionality allows the user to place the node in any arbitrary position.
5. **Drawing a completely new tree** - the tool offers space for drawing a new tree, where on one side of the control there is a list of possible nodes and arrows for connecting nodes, and on the other side there is a space for drawing - canvas. It is possible to transfer nodes from the palette to the drawing space, as well as to connect these nodes with arrows. When nodes are added and names are populated, the tool offers the ability to make a tree structure out of these nodes, as well as to send that tree further to the server to be stored in the database.
6. **Sending a tree to the database** – within the tool there is a service whose methods are called to interact with the server.

The role of the backend application is to enable:

1. **Route for frontend application where a sentence can be analyzed** – when a frontend application sends GET request the backend application forwards this sentence to the syntax analyzer described earlier. This syntax analyzer is written in Python, so it is necessary to call Python script which returns generated syntax trees for given input.
2. **Route for saving the chosen tree in the database** – when an expert reviewed the syntax trees and chose or drew the correct one. This syntax tree is saved along with tags and sentence that has been analyzed.

IV. The Examples Of The Notation Tool

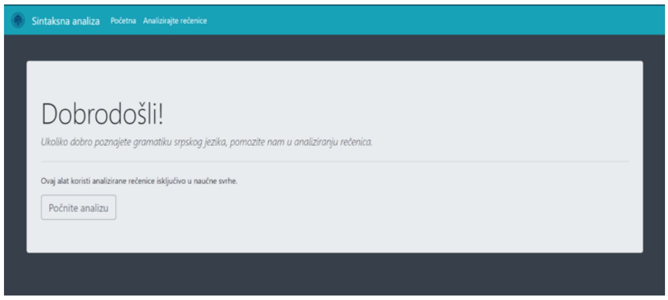


Fig. 4. Welcome page

Figure 4 shows the welcome page. There is a start analysis button that a user can click, and this will open a form for entering the sentence.

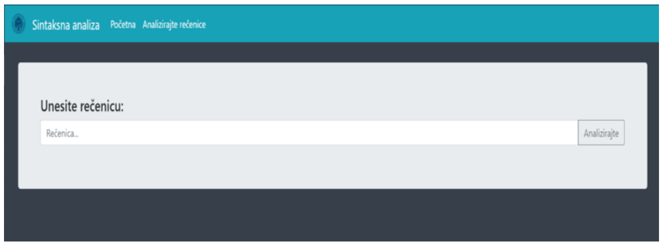


Fig. 5. Enter sentence form

Figure 5 shows a form where the user can enter a sentence for syntax analysis. That sentence is forwarded to the backend application. The backend application sends the sentence to the syntax analyzer by calling Python scripts.

The drawing of syntax trees was implemented using the canvas element in HTML and Canvas API in JavaScript. Below are shown pictures of different options which this tool offers.

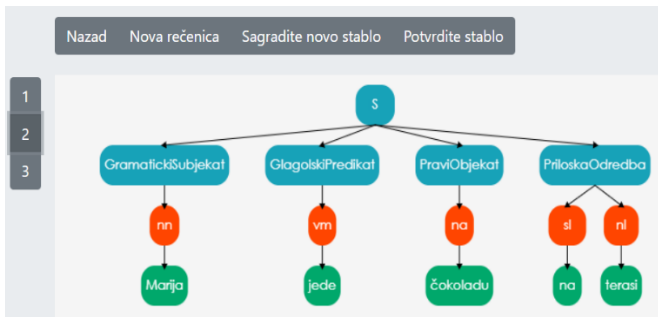


Fig. 6. One of the syntax trees that parser generated

Figure 6 shows the result that parser returned. As can be seen there are three syntax trees generated for this sentence. The second syntax tree is shown in figure 4. The start symbol of the grammar is S, the level below S represents syntactic structures. After that, there are POS tags that tagger returned and finally words of the sentence. The menu above the drawn syntax tree has three options:

1. Interrupting the current analysis and analyzing a new sentence (leads to the form where the sentence is entered),
2. A page where it is possible to build a completely new tree for the current sentence,
3. Confirmation of the current tree - forwarding the selected tree by sending a POST request to the server, where the syntax tree is sent as the request body, after which the syntax tree for the entered sentence is stored in the database.

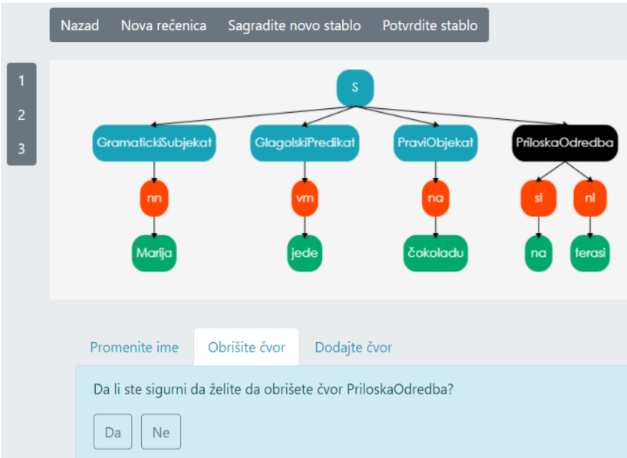


Fig. 7. Node removal

Figure 7 displays the menu for deletion of a node. First, it is necessary to select the node that is going to be altered. The selected node is colored black to stand out from other nodes. After node selection, the application opens the menu where the user can add a new node to the selected one, change the selected node's name or delete the selected node. Figure 7 shows that option for deletion is chosen. If the user wants to rename the node, it is necessary to select the rename option from the displayed menu. After that, the user needs to enter a new node name and confirm it. The third option in the menu is to add a new node. When a node to which a new node is added is selected, it is necessary to enter a name for the new node to be added.

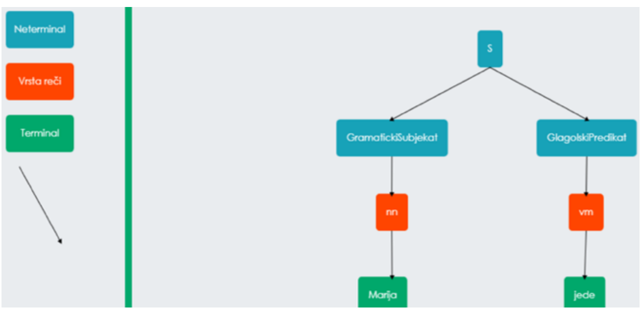


Fig. 8. Drawing syntax tree using canvas

Figure 8 shows the canvas where the user can draw a syntax tree from scratch.

V. Conclusion

The notation tool has been carefully created so that it has the simplest interface with the intention of being used primarily by domain experts - philologists. By using this tool, users are able to tag sentences in the simplest possible way, and thus quickly and efficiently create a corpus for the Serbian language.

After launching this site on the web, it is necessary to hire a set of domain experts who will tag sentences using the tool and create a corpus of sentences. After collecting a sufficient number of sentences, it is expected that these sentences will be used to further improve the Serbian language parser.

Acknowledgment

This work has been supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

References

[1] T. Đorđević, S. Stojković: "Different Approaches in Serbian Language Parsing using Context-free Grammars", Proceedings of 7th International Conference on Electrical, Electronic and Computing Engineering IcETRAN, Etno-Selo Stanišići, Bosnia and Herzegovina (Online conference), pp. 588-591, September 28-30. 2020.
[2] N. W. Francis, H. Kucera, "Brown Corpus Manual" , Technical report, Department of Linguistics, Brown University, Providence, Rhode Island, US, 1979.
[3] M. Davies, *The Corpus of Contemporary American English*, 2008, www.english-corpora.org/coca/, 13.06.2021.
[4] English Corpora, https://www.english-corpora.org, 12.06.2021.
[5] The Quranic Arabic Corpus, https://corpus.quran.com, 10.06.2021.
[6] B. He, B. Dong, Y. Guan, J. Yang, Z. Yang, Q. Yang, J. Cheng, C. Qu, "Building a comprehensive syntactic and semantic corpus of Chinese clinical texts", *Journal of Biomedical Informatics*, vol. 69, pp. 203-217, 2017.
[7] V. Batanović, N. Ljubešić, T. Samardžić, "SETimes.SR – A Reference Training Corpus of Serbian", Conference on Language Technologies & Digital Humanities, Ljubljana, Slovenia, pp. 11-17, 2018.
[8] N. Ljubešić, F. Klubička, "{bs,hr,sr}WaC - Web Corpora of Bosnian, Croatian and Serbian", *Proceedings of the 9th Web as Corpus Workshop (WaC-9)*, Gothenburg, Sveden, pp. 29-35, April, 26. 2014.
[9] MULTEXT-East, https://www.clarin.si/repository/xmlui/handle/11356/1041, 10.06.2021.
[10] ReLDI-NormTagNER-sr 2.1, https://www.clarin.si/repository/xmlui/handle/11356/1240, 05.06.2021.
[11] Syntax tree generator, http://mshang.ca/syntree/, 26.07.2021.
[12] phpSyntaxTree, http://www.tycho.iel.unicamp.br/phpsyntaxtree/?, 26.07.2021.
[13] TreeForm, https://sourceforge.net/projects/treeform/, 26.07.2021.
[14] Trees, https://www.ling.upenn.edu/~kroch/gifdir/Trees3-animation.GIF, 27.07.2021.
[15] Linguistic Tree Constructor, https://ltc.sourceforge.io/screenshots.html, 27.07.2021.
[16] S. Bird, E. Klein, E. Loper, *Natural Language Processing with Python*, Sebastopol, USA, O'Reilly Media, 2009

Modeling the ATP tour matches: A social networks analysis approach

Balša Knežević, Miloš Obradović, Predrag Obradović, and Marko Mišić, *Member, IEEE*

***Abstract*—Professional men's tennis is a demanding sport which greatly benefits from various approaches to performance analysis. More specifically, a complex network theory can be used to model and explain the dynamics of players and tournaments, based on the recorded matches. In this paper, played matches are used to model a social interaction between players. Several undirected weighted networks are constructed to model the ATP tour matches from 2018 to 2020. Moreover, the three most dominant players on the tour (the "Big Three") were observed and analyzed using ego networks approach. The chosen time frame further allowed for the exploration of impact of COVID-19 on the dynamics of the ATP tour. Different network properties were explored, such as small world phenomenon, core-periphery model applicability, community structure, and the rich club phenomenon. Our results based on network theory approach showed that analyzed networks expose similar topological properties, despite the lower numbers of tournaments held in the year 2020.**

***Index Terms*—collaboration network analysis; community detection; ego networks; men's tennis; network modelling.**

I. Introduction

Computational analysis of the results of sports competitions, as well as the performance of teams and individual athletes, has long been present in various sports. The development of data science and artificial intelligence, as well as the possibility of processing large amounts of data, have enabled new approaches to analyze the performance of both teams and individual players. In addition to traditional statistical methods, new methods have been developed, such as collaboration analysis and various prediction techniques.

Several methods based on network science were successfully applied to the analysis of team performance in collective sports, such as football [1][2], basketball [3], and water polo [4]. Furthermore, applications in individual sports are known, such as men's [5][6][7] and women's tennis [8], boxing [9], chess [10], cricket [11], etc. The goal of this paper is to further explore the usage of complex network analysis methodology in the field of men's tennis.

Balša Knežević is with the University of Belgrade - School of Electrical Engineering, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: balsa.knezevic@etf.bg.ac.rs).

Miloš Obradović is with the University of Belgrade - School of Electrical Engineering, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: miobra@etf.bg.ac.rs).

Predrag Obradović is with the University of Belgrade - School of Electrical Engineering, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: predrag.obradovic@etf.bg.ac.rs).

Marko Mišić is with the University of Belgrade - School of Electrical Engineering, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: marko.misic@etf.bg.ac.rs).

The world of tennis tournaments is a complex system that consists mainly of players, the tournaments they play and the matches they have played in those tournaments. Inherently, such a system is very convenient to model with an appropriate collaboration network. Most often, such a system is modeled by players representing the nodes of the network, while the matches that the players play are in some way depicted by the edges in the network.

In this paper, the state of men's professional tennis in the three years from 2018 to 2020 is modeled and analyzed. Similar analyses have already been done in the past for men's tennis in singles [5][7] and doubles [6]. In the meantime, great changes have taken place in the world of tennis. That primarily refers to more than a decade of domination of tennis players from the so-called "Big Three" to which Roger Federer, Rafael Nadal, and Novak Djokovic belong. Moreover, the COVID-19 virus pandemic affected the holding of tournaments in 2020, while tennis tournaments in 2018 and 2019 took place regularly. This allowed for comparative analysis and additional remarks on the impact of the COVID-19 virus pandemic. Therefore, various research methods have been applied in the paper, such as quantitative and qualitative analysis of the collaboration network, community detection, analysis of ego networks of members of the "Big Three", data visualization, etc.

The paper is divided into several sections. The second section describes the studied data sets and provides an overview of the used methodology. The third section presents the results of the research which are then discussed. Appropriate quantitative and qualitative analyses of the data, as well as the produced visualizations, are given. The last section provides guidelines for future work and a brief conclusion.

II. Data sets and methodology of analysis

This section presents the primary dataset and transformations performed on it in order to construct the dataset used for analysis. Furthermore, this section contains the methodology of analysis.

A. Data sets

This paper analyzes the results of men's singles matches played on the ATP tour in the period from 2018 to 2020. Although data from earlier years are available, this timeframe was chosen with the intent to include years 2018 and 2019 which are two consecutive years with regularly held tournament seasons, and the year 2020 which was influenced by the epidemic of COVID-19 disease. Thus, it is possible to determine the influence of pandemics, as the dataset includes both the influenced data points and the two years of regular tennis seasons used as reference points.

The match data was taken on 12/22/2020 from a repository maintained by Jeff Sackman [12] and forms the primary data set for analysis. At the time of analysis, the tennis season for 2020 was completed. The primary data set consists of files containing data on matches in singles competition in the specified period, a list of all players ever ranked on the ATP list, data on the ranking of active tennis players on the ATP list in the period from 2010 to 2020. Match data contains information about the tournament, players, match results with statistics, and the performance of both players during the match. According to the author, the primary data set is largely refined and complete, but there may be certain inconsistencies or incompleteness where the data was not available.

The primary data set contains data on 7117 matches in the specified period. In addition, the data set contains data on 54,975 players who at some point during the observed period had at least 1 point on the ATP list. If several players have the same number of points, then they can share the same ranking on the ATP list, depending on other parameters.

The secondary data set was formed based on the primary dataset, as a refined and cleaned version of it. The data cleaning was performed according to the needs and goals of the research. During the process of cleaning and refinement, some data not necessary for the research itself was intentionally omitted, such as data on players who did not play any matches in the observed period, certain contradictory data, as well as redundant information (columns) that were not used in the analysis. The final secondary data set included data on 7117 matches, as well as data on 581 players.

B. *Methodology of analysis*

Firstly, a thorough statistical analysis of the dataset was conducted. Analyzed properties include the average number of tennis matches in certain years, the average number of tournaments in which tennis players participate, and the ranking of tennis players depending on the number of matches or tournaments played. Most interesting results are presented in the following section. Following the statistical analysis, the refined data set of tennis players and their mutual matches were used to create multiple collaboration networks. These networks were then further studied using methods of complex network theory.

Tennis tournaments are grouped in a season that lasts for a whole year. Therefore, three independent networks were constructed, each holding data for a specific year (N-18, N-19, N-20). Additionally, to allow the analysis of the whole data set, the three networks were aggregated into N-T. The four networks together are referred to as N-series networks.

As per common practice in the field of social network analysis, the network is represented through a set of nodes that describe the actors within the social network and the edges that represent social relations. In the case of networks used in this paper, the nodes of the network are tennis players who played at least one official ATP match in the analyzed period. The two tennis players are connected if they have played at least one official ATP match. The weight of the edge represents the number of matches that the tennis players played with each other. The networks are undirected.

In addition to networks representing all players and matches, the ego networks of the members of the "Big Three" were constructed for each year. These consist of prominent ego nodes, their direct connections with the neighbors, as well as the mutual connections of the neighbors. Furthermore, these three ego networks were unified, and then aggregated They were used to analyze the core-periphery property and the topology of the core of the N-series networks.

Community detection was performed by the Louvain method over the entire network, as well as over the aggregated ego network. For this purpose, a set of filtered and reduced networks was constructed. Clustering strength was evaluated, and the rich club phenomenon was examined.

Python programming language was used to collect and refine data, model the network, and calculate specific metrics using the NetworkX package [13] for network analysis. Gephi [14] was used to visualize and determine network metrics.

III. Results

This section presents the results of the research. The first subsection explores the basic properties of N-series networks, while the rest explores the derived ego networks and community detection.

A. *Basic properties of N-series networks*

A statistical analysis of networks N-18, N-19, N-20, and N-T was conducted. Basic quantifiable features of those networks are presented in Table 1. As expected, the number of tournaments and matches held in 2020 is significantly lower due to the pandemic. This is further reflected in the weighted and unweighted degrees of nodes. However, looking only at the statistical data does not give the whole picture, as it would lead one to believe that year 2020 was significantly different from the previous two years. Only after applying the complex network theory methods discussed below one can give a proper conclusion about the impact of COVID-19 on the dynamics of the observed data sets.

TABLE I
Metrics of constructed N-series networks

	N-18	**N-19**	**N-20**	**N-T**
Players (nodes)	419	364	345	581
Edges	2489	2378	1325	5330
Matches	2974	2696	1447	7117
Tournaments total	138	123	67	328
Tournaments hard surface	81	80	46	207
Tournaments clay surface	47	34	20	101
Tournaments grass surface	10	9	1	20
Avg. weighted degree	13.79	15.28	8.39	24.50
Avg. unweighted degree	11.88	13.07	7.68	18.35
Network density	0.03	0.04	0.02	0.03
Avg. shortest path length	3.13	3.04	3.18	3.23
Diameter	11	9	9	10
Avg. clustering coefficient	0.17	0.19	0.14	0.26

Networks N-18, N-19, and N-20 have an exceptionally low density and a relatively low average shortest path length. Given the low average local clustering coefficient, the networks do not express the small-world property. This is in contrast with previous works in the field [15], but the discrepancies come from a completely different network model. These observations also stand for the aggregated network N-T, as the aggregation does not significantly increase the density nor strengthen the clustering.

Another interesting observation can be made about the average weighted and unweighted degrees. As shown in Table 1, the relative difference between weighted and unweighted degrees is small for networks N-18, N-19, and N-20. This shows that an average pair of tennis players rarely meet more than one time per season. Similarly, in the aggregate network N-T, the annual expected number of matches played by a pair of players is lower than 2. Given the bracket organization of tennis tournaments and loser-go-home policy, only the best players are expected to play multiple matches in a tournament. This leads to the probability of two players meeting in a tournament being quite low, even if they both play in the tournament. In addition, a low annual number of tournaments leads to a low number of annual matches and further decreases the possibility of two players meeting.

A further discussion on this topic can be made when tournament seeding is taken into consideration. The probability of the first and second seed in a tournament gets further artificially lowered, as they are seeded in opposite sides of the bracket and are unable to meet before the finals. If a pair of players is consistently seeded with the top two seeds, this can lead to a measurable decrease in the weight of the edge connecting them.

B. *Analysis of ego networks*

Looking only at the average number of matches played does not show the whole picture and unravel the true topology of the constructed networks. Therefore, a distribution of the number of matches played during the observed period has been calculated and is shown in Fig. 1. As can be seen, many players have only played one or two matches and are thus very isolated, suggesting a core-periphery topology.

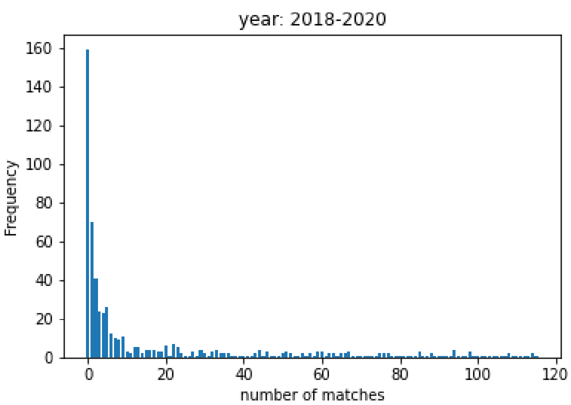


Fig. 1. Distribution of the number of matches played during the three years from 2018 to 2020. The distribution largely resembles a Pareto distribution.

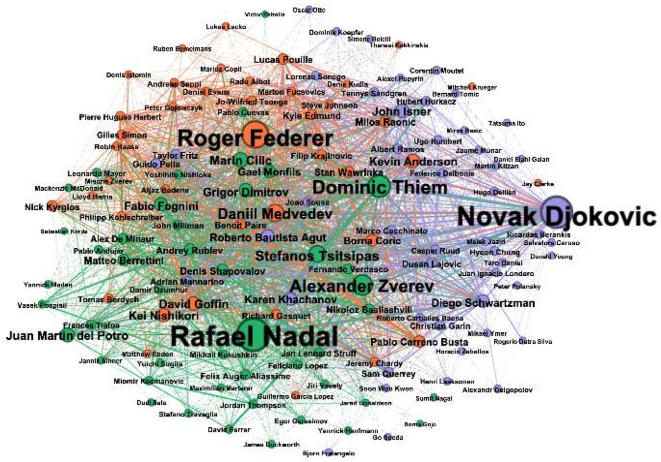


Fig. 2. EGO-T, a unified ego network of the "Big Three" for the period from 2018 to 2020. The size of the node represents weighted node degree and nodes are colored based on clustering.

As stated in the section about methodology, to check if N-18, N-19, N-20, and N-T networks follow the core-periphery model and unravel the topology of the cores of specified networks, several ego networks centered around the members of "Big Three" were constructed. Annual ego networks of Djokovic, Nadal, and Federer were then unified into EGO-18, EGO-19, and EGO-20. Together with these ego networks, their aggregated network EGO-T, shown in Fig. 2, was built.

Clustering the EGO-T network using the Louvain method [16] and tuning the resolution to give 3 clusters reveals a very interesting phenomenon. The original ego nodes bind stronger to some of the other nodes in the network than between themselves. This is in concert with the aforementioned observation about the bracket system and seeding principles influencing the edge weights on the very top of the ATP list.

Exploring the number of nodes and edges of N-series networks included in EGO-series networks can help us explore the properties of the core of N series networks. These statistics are therefore shown in Table 2.

TABLE II
METRICS OF EGO-SERIES NETWORKS

	EGO-18	EGO-19	EGO-20	EGO-T
Nodes	81	88	57	136
Nodes covered	19.33%	24.17%	16.50%	23.4%
Edges	691	744	202	2563
Edges covered	27.76%	31.28%	15.24%	48.08%

Given the percentage of all players and matches included in EGO-series networks, it is obvious that even EGO-T which aggregates other EGO networks and enhances the core property can not be considered a core by itself. Further exploring this topic, the Rombach core finding algorithm [17] was applied to find cores of N-18, N-19, N-20, and N-T, giving cores with 234, 200, 193, and 315 players, respectively. These cores are much larger than EGO series networks and include most of the players.

However, a remark has to be made about the EGO-T network and the percentage of matches included in it. Even though EGO-T is more than two times smaller than the core of N-T, it includes 48.08% of all matches recorded in N-T, which is an astounding amount. This means that matches between the players from EGO-T represent nearly half of all the ATP matches played from 2018 and 2020 and could be used to study some phenomena on a smaller, but representative, group of players, without drastically compromising the number of matches included in the data.

C. Community detection and the rich club phenomenon

To discover a more fine-grained structure in the constructed networks, in addition to exploring network cores, the Louvain method was used once again to find communities in N-18, N-19, N-20, and N-T. Before running the Louvain method, all nodes with degrees lower than 3 were removed from N-18, N-19, and N-20 to avoid the formation of forced and unnatural clusters due to modularity optimization. Characteristics of these reduced networks, aptly named R-18, R-19, and R-20 (R standing for “reduced”), are shown in Table 3. Moreover, a similar procedure was applied to N-T, removing all players with less than 5 matches during the three years, giving us R-T, a reduced network of total aggregated data.

TABLE III
METRICS OF R SERIES (REDUCED) NETWORKS

	R-18	R-19	R-20	R-T
Nodes	244	203	181	287
Nodes retained	58.23%	55.77%	52.46%	49.40%
Edges	2292	2190	1159	4889
Edges retained	92.09%	92.09%	87.47%	91.73%
Communities	9	6	8	7
Avg. clustering coeff.	0.22	0.24	0.18	0.32

The process of node removal is validated by looking at the percentage of nodes and edges retained in the reduced networks. As we can see in Table 3, 49.40% of players played 91.73% of matches during the observed three-year period. This phenomenon can also be seen in Figure 1. As the distribution of the number of matches loosely follows a Pareto distribution, it is to be expected that a rich-club phenomenon can express itself when considering the number of matches as “wealth”. This is somewhat validated by looking at EGO-T, as it consists of a small group of players which bind strongly to each other and monopolize the number of matches over the observed period.

Communities formed by the Louvain modularity clustering are grouped by average rating during the period. This is to be expected, as players of similar ratings choose to play and qualify for the same class of tournaments and are more likely to meet each other. However, the clustering is still not strongly expressed, as can be seen from the average local clustering coefficients.

IV. CONCLUSION

Studying interactions of men’s tennis players proved to be interesting in several aspects. Motivated by the available data, several undirected weighted networks with node metadata were constructed, analyzed, and characterized and multiple common phenomena in the field of complex network theory were explored. Those include small world phenomenon, core-periphery model applicability, community detection, and the rich club phenomenon. In addition, the authors’ own experience with the topic helped explain many of the observed properties and the given explanations are one of the biggest results of this paper, as they give a much better understanding of the dynamics of men’s tennis and are a result of social network analysis and network theory approach to the problem.

In addition, provided network models clearly show an impact of the COVID-19 pandemic on the tennis world, through a smaller number of matches and participants. However, the network theory methodology applied in this paper also shows that the topological properties of the data (such as clustering properties, rich club and small-world phenomena, core-periphery property) stay largely the same, which could not be inhered by naive statistical analysis of the primary data set.

This paper and the constructed networks form a strong basis for further exploration of the topic, including the analysis of mixing patterns in the data depending on the ratings of players, geographical locations of tournaments, affiliations of players, etc. Furthermore, the data in network form is much more suitable for solving some regularly asked questions in the field, such as ranking and match outcome prediction using graph convolutional networks or graph attention models. Lastly, the provided networks are an ideal model for the problem of choosing the representatives of the international tennis community, touching upon the problem of choosing the dominating set of the graph.

ACKNOWLEDGMENT

This work has been partially funded by the Ministry of Education, Science, and Technological Development of the Republic of Serbia. Grant numbers III44009 and TR32047.

REFERENCES

[1] T. Grund, “Network structure and team performance: The case of English Premier League soccer teams”, *Social Networks*, 34(4), pp.682-690, 2012.
[2] J. Gama, M. Couceiro, G. Dias, V. Vaz, “Small-world networks in professional football: conceptual model and data”, *European Journal of Human Movement*, 35, 85-113, 2015.
[3] F. M. Clemente, F. M. L. Martins, D. Kalamaras, R. S. Mendes. “Network analysis in basketball: Inspecting the prominent players using centrality metrics”, *Journal of Physical Education and Sport*, 15(2), 212, 2015.
[4] P. Passos, K. Davids, D. Araújo, N. Paz, J. Minguéns, J. Mendes, “Networks as a novel tool for studying team ball sports as complex social systems”, *Journal of Science and Medicine in Sport*, 14(2), 170-176, 2011.
[5] F. Radicchi, “Who is the best player ever? A complex network analysis of the history of professional tennis”, *PloS one*, 6(2), e17249, 2011.
[6] K. Breznik, “Revealing the best doubles teams and players in tennis history”, *International Journal of Performance Analysis in Sport*, 15(3), 1213-1226, 2015.
[7] U. Michieli, “Complex Network Analysis of Men Single ATP Tennis Matches”, *arXiv preprint arXiv:1804.08138*, 2018.
[8] M. Kostić, D. Drašković, “Complex Network Analysis of Women's Singles Tennis Matches”, Telecommunications Forum (TELFOR), Belgrade, Serbia pp. 1-4, IEEE, 2020.
[9] A. G. Tennant, C. M. Smith, J. E. Chen C, “Who was the greatest of all-time? A historical analysis by a complex network of professional boxing”, *Journal of Complex Networks*, 8(1), cnaa009, 2020.

[10] N. Almeira, A. L. Schaigorodsky, J. I. Perotti, O. V. Billoni, “Structure constrained by metadata in networks of chess players”, *Scientific reports*, 7(1), 1-10, 2017.
[11] S. Mukherjee, “Identifying the greatest team and captain—A complex network approach to cricket matches”, *Physica A: Statistical Mechanics and its Applications*, 391(23), 6066-6076, 2012.
[12] J. Sackmann, Repository tennis_atp, available on: https://github.com/JeffSackmann, accessed: 22.12.2020.
[13] A. A. Hagberg, D. A. Schult, P. J. Swart, “Exploring network structure, dynamics, and function using NetworkX”, Proceedings of the 7th Python in Science Conference (SciPy2008), Pasadena, CA USA, pp. 11–15, August, 2008.
[14] M. Bastian, S. Heymann, M. Jacomy, (2009, March). “Gephi: an open source software for exploring and manipulating networks”, Proceedings of the International AAAI Conference on Web and Social Media, vol. 3, no. 1, 2009.
[15] H. Situngkir, “Small world network of athletes: Graph representation of the world professional tennis player”, Available at SSRN 1001917, 2007.
[16] V. D. Blondel, J. L. Guillaume, R. Lambiotte, E. Lefebvre, “Fast unfolding of communities in large networks”, *Journal of statistical mechanics: theory and experiment*, (10), P10008, 2008.
[17] M. P. Rombach, M. A. Porter, J. H. Fowler, P. J. Mucha, “Core-periphery structure in networks”, *SIAM Journal on Applied mathematics*, *74*(1), 167-190, 2014.

File system performance comparison of native operating system and Docker container-based virtualization

Borislav Đorđević, *Member, IEEE*, Darko Gojak, Nikola Davidović and Valentina Timčenko, *Member, IEEE*

Abstract **- This paper aims to examine and compare the file system capabilities of container virtualization and the native host. Different virtualization categories are mentioned with a focus on OS level types. We have described the importance of container virtualization and its contribution to virtualization popularization. Also, the paper contains a detailed description of the Docker container-based virtualization, its mode of operation, as well as the advantages and disadvantages it possesses. Since the main purpose of this work is to measure the host and Docker file system throughput, one of the best open-source benchmarks is chosen and presented - FileBench, through which all tests were performed. With a practical example, we have shown the file system performance comparisons considering Docker containers and host physical machine.**

Keywords - Docker; containers; virtualization; benchmark; FileBench; file system; performance; comparison.

I. INTRODUCTION

There is rapid development in the IT industry, while hardware and software are changing daily. Hardware development is accompanied by software solutions that aim to make the most efficient use of performance. We strive for solutions that will meet today's standards, asking ourselves what the best use is and how to optimize the available resources so that the requirements and user needs are met.

Some of the most important characteristics in hardware manufacturing are the development costs and time [1]. The above brings us to one of the indispensable topics of today in the IT world - virtualization.

The question is whether virtualization is a better solution and how cost-effective it is, whether it is possible to achieve the desired results with virtualization, and what the limitations are.

There are several varieties of virtualization types, and it can be said for all of these varieties to be usable, with some being more simplified, that is, less decomposed than others. One of the variations is that virtualization can be divided into eight types: hardware virtualization, network virtualization, storage virtualization, memory virtualization, software virtualization, OS level virtualization, data virtualization, and desktop virtualization.

The type of virtualization covered in this paper is "OS level virtualization", whose instances are sometimes called containers. One of the most common associations when mentioning container instances is the well-known Docker [2]. In this paper, the Docker container's file system is examined and compared with the host file system performance.

As the popularity of container virtualization has been growing over time, so have questions about the performance of this type of virtualization. It is hard to talk about container virtualization without mentioning the increasingly prevalent Docker. The ease of installation and use, as well as simplicity of containers management, made Docker a good candidate for file system testing. Another benefit of using it is that Docker containers are lightweight, time savers (it takes less than a minute to build one instance) and besides that, they are consuming a small amount of disk space, so those instances will not affect the host significantly.

Thus, in this paper, the response of the file system of the native operating system and Docker container-based virtualization was researched, and then a comparison of the obtained results was made.

II. RELATED WORK, OBJECTIVE, AND MOTIVATION

As hardware is developing fast today, in terms of storage size, its response speed, as well as processing power, there is an inevitable question about the efficient use of physical machines, which are in most cases underused, or their full potential is not reached [3]. In this regard, scientific research deals with the consideration of further efficiency enhancements possibilities and the mentioned issues.

There is a growing debate about whether virtual solutions are always better and whether they can be expected to largely compete with physical machines [4], [5]. As a big part of the hardware resources in many cases remain unused, there is a lot of room left for the possibility of implementing virtual instances and consideration of the most optimal use.

As the main goal of this paper is to compare the performance of the file systems, with equal settings and the same conditions of the benchmark used for host and Docker

Borislav Đorđević – Institute Mihailo Pupin, Volgina 15, 11000 Belgrade, Serbia, (borislav.djordjevic@pupin.rs)

Darko Gojak – VISER, School of Electrical and Computer Engineering of Applied Studies, Belgrade, Serbia (darkogojak@gmail.com)

Nikola Davidović – University of East Sarajevo, Faculty of Electrical Engineering, Vuka Karadzica 30, 71123 East Sarajevo, RS, BiH, (nikola.davidovic@etf.unssa.rs.ba)

Valentina Timčenko - Institute Mihailo Pupin, School of Electrical Engineering, Belgrade, Serbia (valentina.timcenko@pupin.rs)

containers, we resorted to the method of comparative analysis through FileBench workloads, where four were selected, namely: fileserver, webserver, varmail and randomfileaccess. In our opinion, these are some of the best options for file systems workload testing procedures.

After setting the hypothesis, where it was expected that the physical machine dominates in all fields of given loads comparing to the containers, we proceeded with the application of the experimental method and obtained results that fully justified the assumptions. Based on the comparative analysis method, the obtained results confirmed the initial estimates and expectations, which is proved through the given equations as well as through workloads.

For better understanding and a clearer picture of the container's service capacity, measurements were also performed by increasing the number of Docker instances that worked in parallel, starting from one, until reaching four instances, where all of those were used simultaneously. The decrease of their power was observed and examined.

III. Host OS and Docker

To install the Ubuntu 20.04 operating system on the host, in this case with hardware characteristics shown in Tables 1 and 2, 1024 MiB of RAM is required at least. With Desktop image, which is the most common, there is the ability to try Ubuntu without changing the current computer system. There is also a Server install image that can be only permanently installed on the machine, but without a graphical user interface.

Experienced users are increasingly opting for Ubuntu when it comes to container operations. We can say that the most important item for security, performance, and quality is the Linux kernel, which always has the latest versions of the kernel accompanied by up-to-date security features. All of the above-mentioned is the reason why the world's largest cloud operators opt for Ubuntu operating system to run their containers [6].

Most users will agree and say that Docker became synonymous with container technology, as it had the greatest impact on popularization. But container technology is not a new term, it has been built into Linux in LXC form for more than ten years, and similar virtualization at the operational level systems was offered by: FreeBSD jails, AIX Workload Partitions, and Solaris Containers [7].

Unlike hypervisor virtualization, container virtualization does not have a hypervisor that would be used as a layer of abstraction, isolation of operating systems and applications from the host operating system. There are two types of hypervisors: type 1, which is mounted directly on the hardware, whereas, on the other hand, we can say that the Docker engine is like type 2, which depends on the host operating system, where the Docker container would be in the virtual machine role (Figure 1) [8].

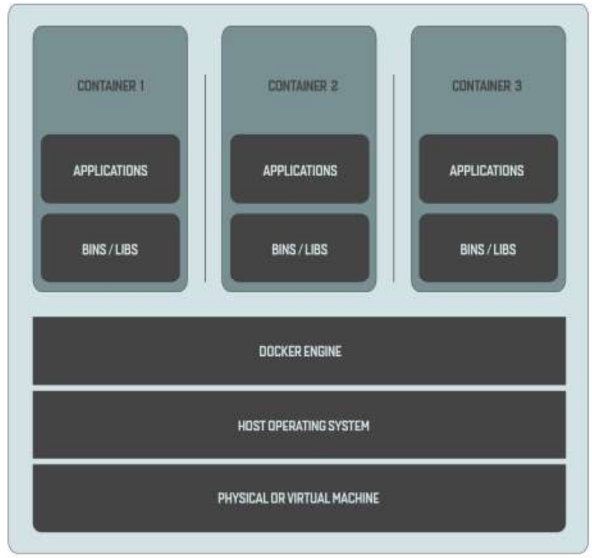


Fig. 1. Docker container-based virtualization

There is a belief that container virtualization is less secure compared with hypervisor virtualization because if weaknesses can be found in the host's kernel on which the containers are located, it could allow intrusion into the containers. The same can be said for the hypervisor, but since the hypervisor provides far less functionality than the Linux kernel (which usually implements file systems, networking, application process controls, etc.) it leaves much less space for attack. In recent years, great efforts have been made to develop software to improve container security. For example, Docker and other container systems now include a signing infrastructure that allows administrators to sign container images to prevent the deployment of unreliable containers [9].

Below is a simple description of docker client-server architecture. Docker client communicates through REST API, over network interface or UNIX sockets with Docker daemon which does building, running and distributing containers (Figure 2) [10]. It is not mandatory that Docker daemon has to run on the same operating system as the Docker client, which can also be connected to a remote daemon [11].

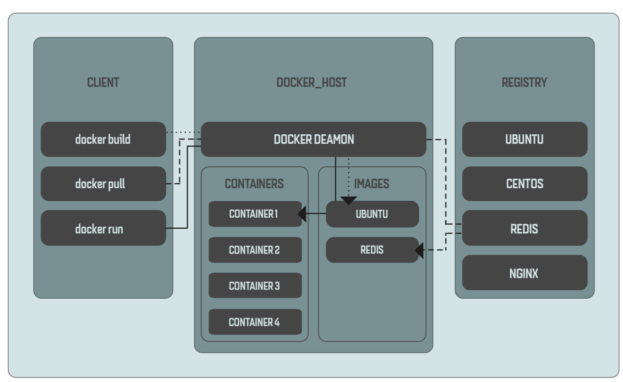


Fig. 2. Docker architecture

Some Linux distributions are designed for running containers and Docker such as Project Atomic [12], Photon OS, RancherOS, etc. [13]. Since 2016, Docker containers have also been able to run on Windows operating system and managed from any Docker client or through Microsoft PowerShell [14].

Docker can also work on popular cloud platforms [15], including Amazon Web Services, Google Compute Engine, Microsoft Azure, Rackspace, etc. [16].

IV. Hypothesis of Expected Behaviour

To make it easier to understand how the results were obtained, the following formulas were derived:

$$T_{WKLD} = T_{RR} + T_{SR} + T_{RW} + T_{SW} \tag{1}$$

In equation 1, the T_{WKLD} notation stands for the total processing time for each workload. This is followed by a random - T_{RR} and a sequential - T_{SR} reading time, while the T_{RW} notation indicates the random write time, and T_{SW} stands for sequential write time. The following formula represents the expected file system access time for each individual workload:

$$T_W = T_{DIR} + T_{META} + T_{FL} + T_{FB} + T_J + T_{HK} \tag{2}$$

The T_W notation above represents the total time required to complete all operations on the ongoing workload. The following notations represent the time required to complete all operations related to: directory - T_{DIR}, metadata - T_{META}, free list - T_{FL}, file block - T_{FB}, journaling - T_{FJ} and house-keeping - T_{HK}. There are two candidates for file system performances that are covered in this paper and they are:

1. native HostOS
2. native HostOS + Docker engine + containers

1. The Ubuntu 20.04 operating system is installed on the host with its default file system, and since the Docker containers are running on it, the native host will play a major role in terms of file system performance. For a better comparison with the host, Ubuntu image is pulled and run on all four container instances. Thus, benchmark and the host file system characteristics depend on the time needed to process benchmark-generated workload, and are noted in the following formula as T_W:

$$T_W(nHost) = f(Bench, hOS\_FS) \tag{3}$$

2. The docker engine has the biggest impact on performance after the host and its file system where everything takes place. As mentioned, HostOS, Docker engine, and containers run on the host file system, except for Docker volumes and self-storage. The benchmark, the host file system characteristics, and Docker engine mapping depend on the time needed to process benchmark-generated workload, in the following formula noted as T_W:

$$T_W(DOCKER) = f(Bench, D\_engine, hOS\_FS) \tag{4}$$

The obtained performance results of the file system of the host and Docker container were predicted by the given formulas. So, as expected, the host was in the lead through all workloads, which was confirmed by the calculation from equation 3. There are small differences in throughput in all segments between the single running container and the host. This lag in the performance of the container was caused by the Docker engine, which was also confirmed by equation 4. After monitoring the throughput of these instances, the following conclusion was made:

Single Docker container is slightly behind the host performances by all measurements, while for any increase of containers running in parallel by one instance, the deterioration in throughput power should be expected.

V. Test Configuration and Benchmark Application

There are various tools, benchmarks that can measure performance in order to examine the capabilities of physical machines as well as the capabilities of virtual solutions. Some benchmark tools are open-source, while others are commercial solutions. Depending on the purpose of the tests, we can opt for one of the most adequate tools. For these measurements, a FileBench is chosen as one of the most suitable benchmarks.

FileBench is a storage and file system benchmark. It uses its own Workload Model Language (WML) that can allow I/O specification of application behavior. It is one of the best-known open-source tools, which, unlike most of the tools that mainly rely on predefined workloads (which cannot be changed in most cases), allows workload modifications as well as adaptation to the specificities of the purpose for which the testing is performed.

Installing a FileBench benchmark is quite simple after downloading the software package. However, on Ubuntu, it requires a few more commands than on Centos operating system, for instance, where it is possible to install it with a simple "yum install filebench" command. Additionally, there is a difference in the installation of the benchmark between two versions covered in this paper. In the first part of the installation, as the configuration files are not included in the repo, they have to be created. Therefore, for the last stable version, it is necessary to run the following commands if they are not installed, respectively: libtoolize, aclocal, autoheader, automake, --add-missing, autoconf.

The second part of the installation requires the installed gcc, flex and bison in order to run FileBench [17]. This part is the same as in the 1.5-alpha3 version, except that in this version it is the only step and it involves running the following commands, respectively: ./configure, make, make install.

In order to measure as accurately as possible and to obtain as better results as possible, Ubuntu 20.04 operating system was installed on the host (hardware shown in Tables 1 and 2) only for this file system test purpose, which after the installation of the benchmark had no other applications that could disrupt the operation of this tool in any way. Also, containers had nothing but installed FileBench.

After everything is set, there is still one thing left to do and that is disabling ASLR (address space layout randomization) by changing the value of randomize_va_space to “0” (zero), otherwise, the workloads will be blocked in the stage of running.

TABLE I
HARDWARE CONFIGURATION OF THE HOST

Component	Characteristics
Processor	AMD Ryzen 5 3600X, 3.8GHz – 4.4GHz, 6 Core, 12 Thread
Cache	L1 Cache 384KB, L2 Cache 3MB, L3 Cache 32MB
Memory	16Gb DDR4, 3200MHz
SSD	Kingston A2000 SA2000M8/500GB
Motherboard	GIGABYTE B450M DS3H

TABLE II
SSD characteristics

Capacity	500GB
DRAM	DDR4
Interface	NVMe™ PCIe Gen 3.0 x 4 Lanes
Form factor	M.2 2280
NAND	3D TLC
Sequential Read/Write	up to 2.200/2.000MB/s
Random 4K Read/Write	up to 180.000/200.000 IOPS

VI. TESTS AND RESULTS

Each measurement was done in three rounds per host and per each container instance, after which the average value was taken for results. The obtained measurements of individual container performances were then compared with the results obtained while testing the host. The throughput of each container was observed in cases when only one container instance was started, when two instances were running in parallel, and when three and then four containers were running at the same time.

File system performance tests were conducted on the latest stable version of FileBench - 1.4.9.1 and 1.5-alpha3 version where throughput was measured in MB/s. For the purposes of this experiment, four of the over fifty predefined workloads were selected. On both versions, the performance of the filesystem was tested via three workloads that were used to emulate applications, namely: fileserver, webserver and varmail. On the last stable version, an additional workload was included - radnomfileaccess. The following is a brief description of workloads that were used and covered with formulas (1) and (2): *Fileserver* – It mimics the elementary I/O activity of a file server. It performs a sequence of creating, deleting, adding, reading, writing, and attribute operations on a directory tree; *Webserver* - Mimics elementary I/O activity of a web server. Produces an open-read-close sequence on multiple files in a directory tree, plus appends a log file; *Varmail* - Imitates elementary I/O activity of a mail server that saves each e-mail in an isolated file (/var/ mail/server). It contains a set of multiple threads of the following operations in a particular directory: create-add-sync, read-add-sync, read, and delete; *Randomfileaccess* - Uses random variables that are user-defined entities, and these entities are formulated by a random distribution that is used to select a random value that is returned with each use [18].

It is hard not to mention virtual clusters when Docker containers are used. Testing could take on a completely different dimension if any container orchestration platforms such as Kubernetes were used, where containers would combine and pool their serving powers [19]. But the purpose of these tests was to compare the file system performance of the host and individual container.

The parameters shown in Tables 3 and 5 are set with default values. The values for the four specified parameters (number of files - nfiles, average file width, and size - (meandirwidth, meanfilesize), as well as the number of threads - nthreads) are the same in both versions of the benchmark. The time for executing each of the workloads is set to 60 seconds, which is the default value for most of the predefined workloads.

TABLE III
PARAMETERS OF THE SOURCE CODE *.F FILES (1.4.9.1 VERSION)

Workload (runtime 60s)	Fileserver	Webserver	Varmail	RFA
nfiles	10.000	1.000	1.000	10.000
meandirwidth	20	20	1.000.000	20
meanfilesize	128k	16k	16k	Random
nthreads	50	100	16	5

TABLE IV
BENCHMARK RESULTS (MB/S), 1.4.9.1 VERSION

Instance	Fileserver	Webserver	Varmail	RFA
Host	3866.6	1001.5	187.4	19081.8
1 container	3746.1	962.8	180	18190.1
2 containers	1764.4	695.9	158.3	9438.5
3 containers	1170.5	528.5	137.2	5300.7
4 containers	651.7	458.8	117.2	3809.8

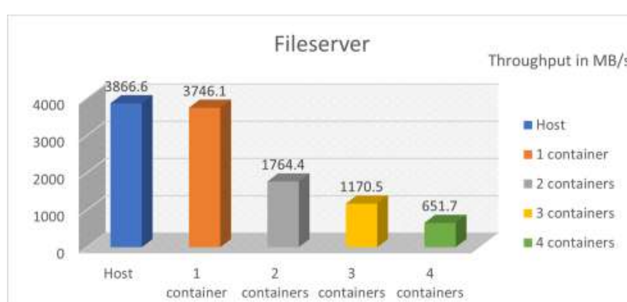


Fig. 3. Fileserver test results from Table 4

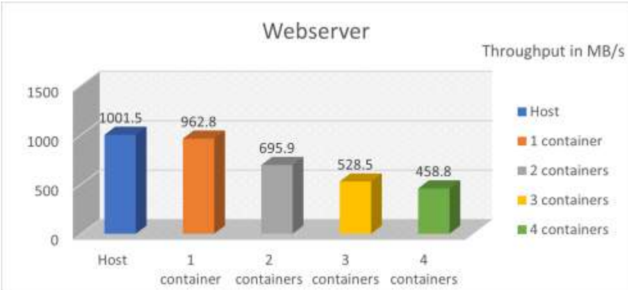


Fig. 4. Webserver test results from Table 4

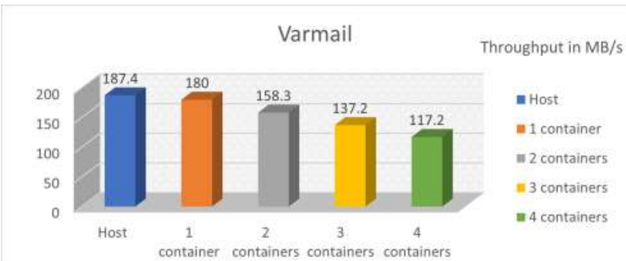


Fig. 5. Varmail test results from Table 4

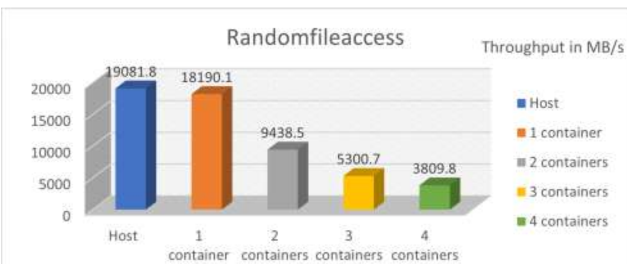


Fig. 6. Randomfileaccess test results from Table 4

A. Measurements performed on version 1.4.9.1

The host had better performance in all four categories which is shown in Table 4. The obtained results were proved by formulas (3) and (4). Starting with the fileserver environment, there is a small throughput difference of 3 % in favor of the host compared to a single container. Then, as expected, by increasing the number of containers by one, the serviceability also decreases, so that the performance of the two running containers drops by more than twice, i.e. 54%. Performance with three running containers deteriorated by 70% and with four instances the results showed it to be 83% (Figure 3).

For webserver tests, the results are as follows. The throughput at the host instance is 4% higher when compared to a single running container, while for two running containers that gap is 30%. With three and four containers in running state, we can see the degradation of 47% and 54%, respectively (Figure 4).

In the case of varmail environment, the single running container has lower performances by 4%, two containers by 16%, and three and four containers by 27% and 37% compared to the host (Figure 5).

The randomfileaccess workload also had poorer container results, showing performance declines of 5, 51, 72, and 80% when having 1, 2, 3, and 4 containers in running state, respectively (Figure 6).

TABLE V

PARAMETERS OF THE SOURCE CODE *.F FILES (1.5-ALPHA3 VERSION)

Workload (runtime 60s)	Fileserver	Webserver	Varmail
nfiles	10.000	1.000	1.000
meandirwidth	20	20	1.000.000
meanfilesize	128k	16k	16k
nthreads	50	100	16

TABLE VI

BENCHMARK RESULTS (MB/S), 1.5-ALPHA3 VERSION

Instance	Fileserver	Webserver	Varmail
Host	4072.6	3333.8	163.5
1 container	4007.1	3080.1	133.4
2 containers	1696.7	1563.5	111.2
3 containers	704.6	1160.8	88.5
4 containers	451.4	952.6	76

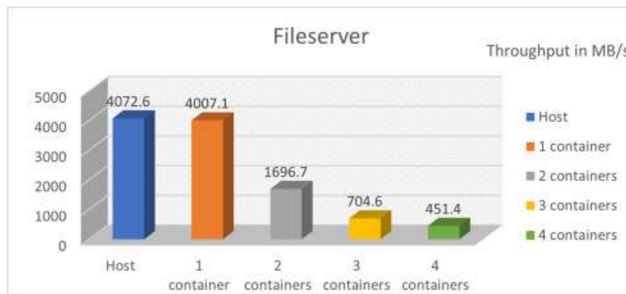


Fig. 7. Fileserver test results from Table 6

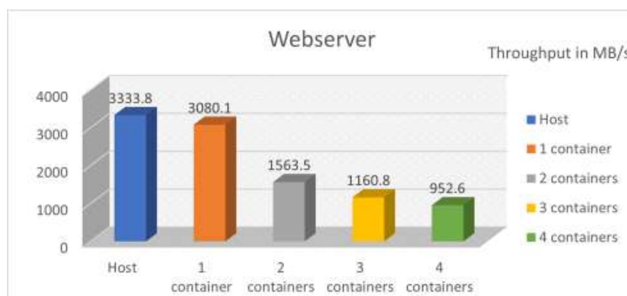


Fig. 8. Webserver test results from Table 6

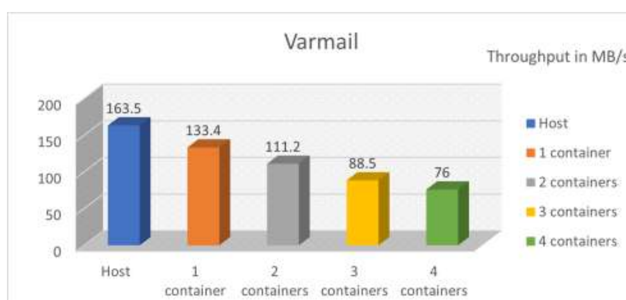


Fig. 9. Varmail test results from Table 6

B. Measurements performed on version 1.5-alpha3

Within a FileBench version 1.5-alpha3, the expected results were obtained, which is verified by formulas (3) and (4). As well as in the latest stable version the host dominates (Table 6). In the fileserver case, a single container performance does not significantly differ from the host and it is lower by 2%, while for two container instances in running state the drop is much bigger, 58%. For three and four instances it is 83% and 89%, respectively per container (Figure 7).

With webserver workload tests we have a throughput deterioration comparing to host, namely 8% for a single container, 53% in the case of two instances, 65% for three running containers, and 71% per instance in the case of four containers running (Figure 8).

As for varmail, the host throughput is higher by 18% compared to a single container, while for two instances there is gap of 32% per instance, it is 46% for three instances and 54% for all four containers (Figure 9).

VII. Conclusion

According to the shown tests, the host had better performance in all segments compared to Docker containers which justifies the hypothesis. During performance monitoring through all four workloads, a slight differences in throughput between the host and single container is noticeable. As we can see in the obtained measurement results, the increase of the number of container instances decreases their service power, which also differs from workload to workload. Those are expected results, and accordingly, depending on the load, we can determine whether containers are suitable and if they will meet the requirements for which container instances were originally intended.

This is only a small segment in testing the host and Docker container capabilities, as there are over forty predefined tests left, as well as many variations of modifying existing and writing your own workloads that can be processed. Since FileBench workloads can be easily managed it leaves a lot of room for future measurements and comparisons with the results of other benchmarks that are not so flexible in terms of tests.

Today, it is known that hardware development is increasingly focusing on multi-core solutions that can process many instructions in a very short time. That leaves plenty of room for further processing of power and resources, which is suitable for the normal and smooth operation of virtual solutions. Virtualization is not always the answer to everything, for some purposes virtualization simply does not achieve the desired results so in that case, the only choice is a physical machine. But in most cases, security, productivity, and cost-reducing benefits outweigh all problems, and therefore Docker virtual solutions and virtualization, in general, are increasingly gaining in popularity.

Acknowledgment

The work presented in this paper has partially been funded by the Ministry of Education, Science, and Technological Development of the Republic of Serbia.

Literature

[1] C. Walls, “Hardware and software development; what’s the cost?,” 2018 [online]: https://www.embeddedcomputing.com/technology/software-and-os/hardware-and-software-development-what-s-the-cost
[2] IBM Cloud Team, IBM Cloud. Containers vs. VMs: What’s the difference? IBM, 2020. [online]: https://www.ibm.com/cloud/blog/containers-vs-vms
[3] Spiceworks. The 2020 State of Virtualization Technology, 2019. [online]: https://www.spiceworks.com/marketing/reports/state-of-virtualization/
[4] K. Thompson, “Hardware vs. Software development: Similarities and Differences,” Cprime, 2015. [online]: https://www.cprime.com/resources/blog/hardware-vs-software-development-similarities-and-differences/
[5] T. Collins, “Virtual servers vs physical servers: Which is best? 10 March,” Atlantech, 2020. [online]: https://www.atlantech.net/blog/virtual-servers-vs-physical-servers-which-is-best
[6] Canonical. Why is Ubuntu #1 OS for containers? Ubuntu, 2018. [online]: https://ubuntu.com/containers
[7] S. Hogg, “Software Containers: Used More Frequently than Most Realize,”, Networkworkd, 2014 [online]: https://www.networkworld.com/article/2226996/software-containers--used-more-frequently-than-most-realize.html
[8] U. Hiwarale, “Anatomy of Docker,” [online]. 2018 Nov [Accessed 24 February 2021]. Available from: https://itnext.io/getting-started-with-docker-1-b4dc83e64389
[9] P. Rubens, “What are containers and why do we need them?,”, Cio, 2017 [online]: https://www.cio.com/article/2924995/what-are-containers-and-why-do-you-need-them.html
[10] N. Poulton, Docker Deep Dive, JJNP Consulting Limited, Lean Publishing. Leanpub book, 2018.
[11] Docker. Docker overview. [online]: https://docs.docker.com/get-started/overview/#docker-architecture
[12] SP. Kane, K. Matthias, “Atomic hosts,” in Docker: Up and Running. 2nd ed. O’Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472; 2018.
[13] JM. Scheuermann, “A Comparison of Minimalistic Docker Operating Systems,” Inovex, 2015. [online]: https://www.inovex.de/blog/docker-a-comparison-of-minimalistic-operating-systems/
[14] M. Friis, “Build and run your first Docker Windows Server container,” Docker, 2016. [online]: https://www.docker.com/blog/build-your-first-docker-windows-server-container/
[15] C. Ward, “The Shortlist of Docker Hosting,” CloudBees, 2016. [online]: https://www.cloudbees.com/blog/the-shortlist-of-docker-hosting/
[16] J. Turnbull, “Installing Docker,” In The Docker Book. CC BY-NC-ND 3.0; 2018.
[17] G. Amvrosiadis, “FileBench,” GitHub, 2016 [online]: https://github.com/filebench/filebench
[18] V. Tarasov, “Predefined personalities. [online]. 2016 Jul [Accessed 24 January 2021]. Available from: https://github.com/filebench/filebench/wiki/Predefined-personalities
[19] IM. Aidan, H. Sayers, “Using a Kubernetes cluster,” in Docker in Practice. Manning Publications Co. 20 Baldwin Road PO Box 761 Shelter Island, NY 11964; 2016.

Performance comparison of native host vs. ESXi hypervisor-based virtualization

Borislav Đorđević, *Member, IEEE*, Srđan Milenković*,* Nikola Davidović and Valentina Timčenko, *Member, IEEE*

Abstract – The main objective of this paper is performance comparison of hypervisor-based virtualization with VMware ESXi virtual machines and native host machine. From all performance classes, for the needs of this research we have chosen the evaluation of the file system performance. The measurements are carried out under equivalent conditions and by a unique test method, using the Filebench software, which guarantees equality and independence from the impact of hardware and operating system characteristics. As the base operating system we have used CentOS 7.7 with the latest updates, while ESXi 6.7 was used as the hypervisor. Performances are compared for the native host machine and ESXi server with one, two and three virtual machines (VM) running simultaneously. We have also analysed the expected behaviours, verified the assumption with Filebench testing software, and provided the concluding remarks for this papers research topic.

Key words – Virtualization; Filebench; Hypervisors; ESXi; VMware; CentOS; Virtual Machines

I. INTRODUCTION

In IT world, the term virtualization refers to the act of creating a virtual version of something, or it is the process of creating and running a virtual instance of a computer resource in a layer abstracted from the actual hardware. It is used to describe virtual computer hardware platforms, storage devices, network resources, server infrastructure, etc. We can experience virtualization in almost all segments of today's computer technology. The main idea behind virtualization is a very simple and came from the corporative approach: the need to satisfy the increase in the utilization of available hardware resources, while at the same time reducing the costs of the infrastructure. Virtualization did exist as a technology even some 30 years ago, but the hardware of those days could not exploit the full usage that virtualization brought, so it was disregarded until progress was made in computer technology giving to virtualization a new meaning, shaping it to what it looks today. Nowadays, thanks to this technology it is possible to run multiple independent operating systems on one physical server. Some of the benefits that virtualization provides are primarily related to saving the necessary physical space that would be needed for the accommodation of the devices and also the electrical energy consumption that would inevitably be used for powering such devices. Today, the use of virtualization in a simple way increases server availability and isolation, making it one of main reasons why these technologies are so popular [1]. When using these technologies it is important to mention that the level of hardware utilization of servers without virtualization is in the range of 15% of its maximum capacity, while with the use of virtualization technologies the utilization raises to more than 70%. These technologies however come with a price, or to be exact, with the retention or even increasing availability of resources, while it is realistic to expect a somewhat lower performance of virtualized systems when compared to the non-virtualized systems, which is the main topic of this paper.

There are several virtualization types: virtualization of hardware, software, desktop, data, network, memory, storage, etc. We are focused on hardware virtualization. Hardware virtualization implies the use of a hypervisor, a layer that acts as a mediator between the host and virtual machine, which is nothing more than a simulated computing environment that can, but does not have to be equal to the physical environment that it simulates. In addition to the classification by the location of the hypervisor layer, the hardware virtualization also depends on what type of virtualization is provided, and can be categorized as: full, hardware-assisted, and paravirtualization.

Full (native) virtualization is a virtualization technique that completely simulates the underlying hardware. Hardware-assisted virtualization (Intel VT-x or AMD-V) is platform virtualization approach that enables efficient full virtualization using help from hardware capabilities, primarily from host processors. In this situation, the processor simulates hardware that does not have to be the same as physical. Paravirtualization is an enhancement of virtualization technology in which a guest operating system is modified prior to the installation inside a virtual machine in order to allow all guest OS within the system to share resources and successfully collaborate, rather than attempt to emulate an entire hardware environment [2].

The remainder of this paper will be structured as follows. Section II provides a brief description of the technologies that are mentioned in the paper and a short review of related work for this project. Section III provides the description of the

Borislav Đorđević – Institute Mihailo Pupin, Volgina 15, 11000 Belgrade, Serbia, (borislav.djordjevic@pupin.rs)
Srđan Milenković - School of Electrical and Computer Engineering of Applied Studies, Vojvode Stepe 283, 11000 Belgrade, Serbia, (smilenkovic1992@gmail.com)
Nikola Davidović – University of East Sarajevo, Faculty of Electrical Engineering, Vuka Karadzica 30, 71123 East Sarajevo, RS, Bosnia and Herzegovina, (nikola.davidovic@etf.ues.rs.ba)
Valentina Timčenko - Institute Mihailo Pupin, School of Electrical Engineering, Belgrade, Serbia, (valentina.timcenko@pupin.rs)

performance-measuring tool that we have used for this experiment. In Section IV, we present a short description of the architecture of the used hypervisor. Section V presents the hypothesis and methodology used to achieve performance comparison. In Section VI, we present the test environment and configuration for the experiment. Test results for our benchmarks' tests are presented in Section VII. In Section VIII, we draw conclusions to the work made in this paper.

II. RELATED WORK AND OBJECTIVE

This paper is primarily devoted to analysis of the performances of hypervisor-based virtualization with one of the most commonly used hypervisors. The hypervisor serves as a layer between the virtual machine's operating system and the host's physical memory, providing data integrity and isolation of VMs. Thanks to hardware-assisted virtualization which is accomplished via EPTS (extended page tables, for Intel chipsets) or RVI (rapid virtualization indexing, for AMD) we have a large increase of speed compared to software memory virtualization [3]. The paper considers advantages of using virtual machines while creating modern network infrastructure; as well as describes an experiment using common test environments and programs for measuring and analysis of hypervisors and their performances. Benchmarking is a popular approach nowadays for many devices and general I/O performance analysis, whereas the special attention is put on the problem of fast input/output support [4-6].

Main contribution of this paper is the examination of the performances of the native host operating system and hypervisor-based virtualization of VMware ESXi [7] [8]. As the technology that was used for this research is relatively new, there are not many references in literature that research with similar environments, tools, and test characteristics. The goal of this paper is to examine the file system performance of the generated workload through Filebench software tool for: (1) mail server scenario which is dominated by random read and random write components; (2) web server scenario where random read components dominate; (3) file server scenario in which both random and sequential components are equally represented; and (4) random file access scenario dominated by random read component [9].We have set up a model for the file system performance analysis of the native host and ESXi based virtual machines. The results of this experiment should give us a full picture of how the performance of a native machine compares to the performance of a hypervisor-based virtual machine.

III. FILEBENCH

Filebench is a software test environment (usually called a benchmark) used to measure the performance of various parts of an operating system. What sets Filebench apart from other benchmarks is the fact that it is equipped with several predefined workloads, which allows users to easily test their systems in various forms (most popular forms being a mail server or a file server) [10]. Presently many benchmarks hard code the workloads they generate quite rigidly, meaning that a user can specify some of the basic workload parameters, but cannot really control the execution flow of the workload in detail. Filebench gives its users freedom to define workloads using a Workload Model Language (WML). WML is mainly composed of four main parts: fileset, process, thread, and flowop.

A standard Filebench test is executed in two stages: fileset pre-allocation and a workload execution. First part of any workload execution is defining a fileset that it uses. A fileset is a named collection of files and to define it a user must specify its name, path, number of files, and a few other optional attributes that can be included in a filesets creation. After defining a fileset the next step are the processes in WML that represent real UNIX processes which are created by Filebench during the test. Every process is made of one or more threads representing an actual POSIX threads and every thread executes a loop of flowops. A single flowop is a representation of a file system operation that is translated to a system call by Filebench.

The ending of a WML file usually contains one of two "run" commands (run and psrun) that tell Filebench to allocate the defined filesets, prepare the required number of UNIX processes and threads, and start a cycled flowops execution. After completing a run, Filebench gives a number of different metrics, where the most important one for the user is operations per second. This is the total number of executed flowop instances (in all processes and threads) divided by the time it took for a full run of the workload. To generate a workload and start the measurement of a particular part of the system, one must execute the *filebench -f workload.f* command.

IV. ESXi HYPERVISOR

VMware ESXi (Elastic Sky X "integrated") is a type-1 hypervisor developed by VMware for deploying and serving virtual machines that was made from its predecessor ESX. Type-1 hypervisors run directly on the host's hardware to control the given hardware and to manage guest operating systems, and for this reason they are mainly called bare metal hypervisors (Figure 1). A guest operating system runs on another level above the hypervisor.

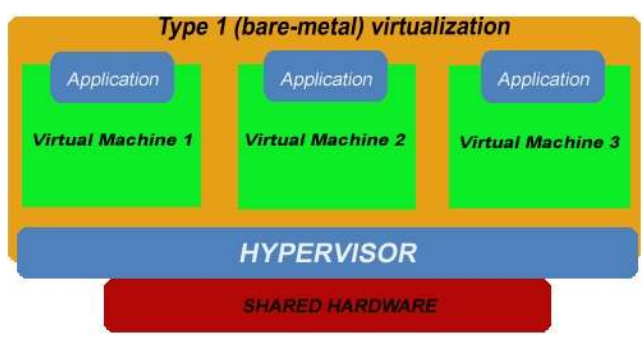


Fig. 1.Type-1 (bare metal) hypervisor

VMware ESXi is a hypervisor that runs on the host server hardware without the underlying operating system. ESXi provides a virtualization layer that abstracts the CPU, storage, memory and networking resources of the physical host into multiple virtual machines. That means that applications running in virtual machines can access these resources without direct access to the underlying hardware. VMware refers to the hypervisor used by VMware ESXi as VMkernel and it receives requests from virtual machines (as processes that run on top of it) for resources and presents the requests to the physical hardware [12-14]. The kernel also provides means for running all processes on the system, including management applications and agents as well as virtual machines. It has control of all hardware devices on the server, and manages resources for the applications as shown in Figure 2 [15]. The main processes that run on top of VMkernel are:

- Direct Console User Interface (DCUI) — the low-level configuration and management interface, accessible through the console of the server, used primarily for initial basic configuration.
- The VMM, virtual machine monitor, which is the process that provides the execution environment for a virtual machine, as well as a helper process known as VMX. Each running virtual machine has its own VMM and VMX process.
- Various agents are used to run and enable high-level VMware Infrastructure management from remote applications.
- The Common Information Model (CIM) system is the interface that enables hardware-level management from remote applications via a set of standard APIs.

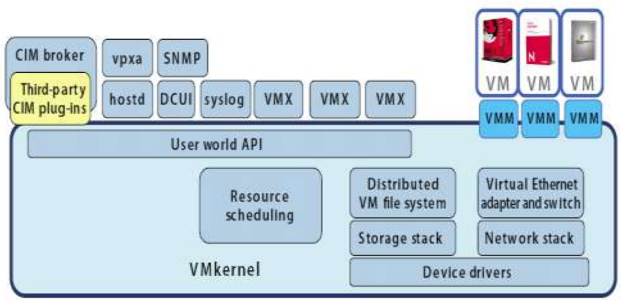


Fig.2.VMware ESXi architecture

V. Hypothesis of Expected Behaviour

Since we are using a Type-1 hypervisor that works directly on hardware, the total processing time for each workload T_W can be described by the following equation:

$$T_W = T_{RR} + T_{RS} + T_{WR} + T_{WS} \tag{1}$$

where T_{RS} and T_{RR} represent sequential and random read time respectively, while T_{WR} and T_{WS} represent random and sequential write time respectively. For every specific workload we have an expected access time for the file system which includes five components as shown in following equation:

$$T_{WORKLOAD} = T_D+T_M + T_{FL}+T_{FB}+T_J +T_{HK} \tag{2}$$

where $T_{WORKLOAD}$ represents the overall time for finishing all operations on the current workload, and T_D, T_M, T_{FL}, T_{FB}, T_J, T_{HK} represent time needed for completing all operations related to directory, metadata, free list, file block, journaling and house-keeping operations in the file system, respectively.

In this study we have a specific situation where there are two sides which have identical settings of the operating system (CentOS) and the file system (XFS), used in the performance testing: (1) Native machine (hostOS) and (2) ESXi + VMs(guestOS).

1. Native hostOS: The time to process the generated workload depends on the benchmark interaction with the hostOS file system and also the characteristic of the file system. Total time to process the workload, $T_{W(native)}$ is defined as:

$$T_{W(native)}=f(benchmark,\ hostOS\_FS) \tag{3}$$

2. ESXi + VMs(guestOS): The time to process the generated workload ($T_{W(ESXi)}$) in this case depends on the benchmark interaction with guestOS file system, the characteristic of the file system and the virtualization processing component of the ESXi hypervisor (ESXi_proc) is as in the following formula:

$$T_{W(ESXi)}=f(benchmark,\ guestOS\_FS,\ ESXi\_proc) \tag{4}$$

Since we use the same settings as the native machine for our virtual machines, the benchmark interaction and characteristics of the file system on the guest will be the same as the ones on the native machine. The virtualization processing component depends on the virtualization type and hypervisor processing as in the following formula:

$$ESXi\_proc = f\,(virt\_type,\ hyp\_proc) \tag{5}$$

In the context of virtualization type, ESXi uses full virtualization, which is further enhanced with one of the technologies (depending on hosts' CPU) for hardware assisted virtualization. In the context of the hypervisor processing it is important to consider the delay, which represents the time required for the hypervisor to receive requests from virtual hardware of a guest OS and forward them to the hosts' hardware for proccessing. The delay can be explained as following: virtual machines generate workload, which passes from a VM through the hypervisor onto the hosts' hardware. First the benchmark application generates the workload which is passed on for further processing to the hypervisor. The second part happens inside the hypervisor and is defined as the interaction between guest workload and VM image file. Generated workload is passed on the hypervisor, which maps it into requests for VM large image files. Lastly the hypervisor's mapping process generates input files as requests for real disk drivers on the hosts' hardware. The time needed for generating those requests depends on the hypervisor's file system and caching capabilities.

The expected outcome according to formula (3) is that the native host will perform better than ours ESXi virtual machines. Virtual machines have a complex data path, formula (5), where data must pass through guest OS file system and the hypervisor onto machine hardware. Therefore, it is expected that a degradation of the ESXi VM performance will happen compared to the native host machine, formula (4).

We have investigated a few cases for the this paper: firstly, the performace of a native host machine, then the performance of a single ESXi VM running and lastly the performance of several virtual machine running at the same time. In general, we expect:

- Native host to perform better when compared to ESXi with one virtual machine running.
- Running several instances of the ESXi virtual machines, *n**ESXi VMs (*n=1,2,3…*), should have a significant performance degradation compared to the native host.

VI. TEST ENVIRONMENT CONFIGURATION

The assumption of an adequate testing is the application of a single hardware configuration, the same operating system, and measurement methodology for all test procedures as mentioned before. The hardware configuration contains all the components necessary for a modern-day computer, and in this case, it is a home-based system of the newer generation (Table 1). CentOS version 7.7 is selected as the operating system, which is currently one of the most popular Linux distribution.

During the installation process we opted for Gnome graphical interface installation option with essential packages and programs for a graphical environment. The XFS file system characteristics and layouts are shown in Table 2. Filebench is a program designed to measure the performance of file systems and storages, and it is capable of generating multiple workload types that simulate environments when using certain servers/services such as mail, web, file, database, etc. Before starting tests, we made sure that all available updates were installed. Each virtual machine was given 4 GB of RAM.

TABLE I
HARDWARE CONFIGURATION OF THE TEST PC

MB	Gigabyte B75M-D2V
RAM	DDR3 1330 MHz, 16 GB
CPU	Intel
Model	Pentium G860
Cores	2 /2 threads
Speed	3.00 GHz
Cache(L1,L2,L3)	2x32kB; 2x256kB, 3MB
SSD	Samsung SSD 860 EVO
Interface	SATA 6Gbps
Capacity	250 GB
OS	CentOS 7.7.1908.el7

This benchmark behaviour is controlled using files with the extension *.f that are written in Workload Model Language, that can be edited in any text editor. The use for individual measurements involves putting a command from a terminal with root privileges using the name of the *.f file as an argument.

TABLE II
FS LAYOUT

FILE SYSTEM	SIZE	MOUNT
/DEV/MAPPER/DATA-ROOT	35 GB	/
/DEV/SDA1	4 GB	SWAP
/DEV/SDA2	1024 MB	*/ BOOT*

VII. TESTS AND RESULTS

The focus of this paper was to measure the performance of hard disks and data-flow in one of the more popular virtualization systems, especially in cases where several instances of virtual machines are being used. The main idea was: as the number of instances increases, there is a significant drop in performance and this drop is constant on any hardware-software configuration. Benchmark of the host computer without virtualization was taken as a reference point for file system performance in these tests.

A number of modified files of the source code ***fileserver.f***, ***webserver. f***, ***randomfileaccess.f*** and ***varmail.f*** were used during the tests, which are thus testing the files, web and the mail server environments, respectively. The changes were taken into consideration when setting the benchmark parameters in a way to provide as realistic as possible exploitation conditions. And while the location (/ bench), the I/O block size (iosize = 1M) and the average size of the add-on (meanappendsize = 16k) are common denominator for all tests, the parameters such as the number of files (nfiles), the average depth of the directory (meandirwidth) the average file size (meanfilesize), cache and the number of threads (nthreads) are changed on a case-by-case basis (with * .f files). The defined settings are retained throughout the entire benchmark test and are displayed in Table 3. For an easier view in the following table the name of each benchmark workload has been abbreviated with their initials (file server (FS), web server (WS), mail server (VMail) and random file access (RFA).

TABLE III
SETTINGS OF THE SOURCE CODE IN THE *.F FILES

	FS	WS	VMail	RFA
nfiles	10.000	1.000	1.000	10.000
meandirwidth	20	20	1.000.000	20
meanfilesize	16k	16k	16k	
nthreads	50	100	16	5
cached				false

The duration of each test was 120 seconds, which is also stated in the *.f files, with the goal of acquiring the most realistic results. Special attention was paid to keep the OS clean and the impact of any external subject on system components was reduced to the minimum. After performing a reference measurement of the host computer without virtualization, ESXi was installed and three virtual machines were generated. Tests were conducted in a way that one virtual machine was first started and measured, then two and three machines simultaneously. From the generated data, the final conclusions were made by calculating the average values of the results.

TABLE IV
BENCHMARK RESULTS (MB/S)

	FS	WS	VMail	RFA
Host	401.6	127.9	51.4	7379.5
1VM	230.4	67.6	45.7	3646.0
2VM	122.3	42.8	24.4	2126.9
3VM	78.2	24.9	14.8	1498.6

Table 4 shows the data we collected from workloads running in the test environment (again we used the same abbreviations like in Table 3). Data from Table 4 are shown on the next few figures, with remarks on the performance displayed in each. All of the measures shown in the following figures are displayed in megabytes per second (MB/s).

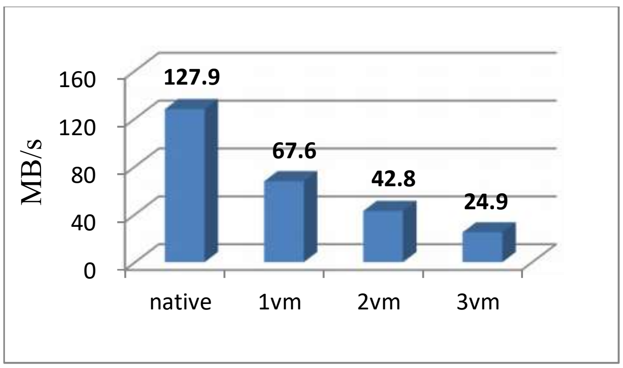


Fig. 3. Webserver.f workload test results

The characteristics of the webserver.f workload with our specification (100 threads) is that random reads dominate, there are some random write components, while the sequential components are not present. Here we observe that native host OS performs much better than in the case with one instance of the ESXi virtual machine (Figure 3). In the case of this workload, instantiation of more than one ESXi virtual machine brings some performance degradation but not significant.

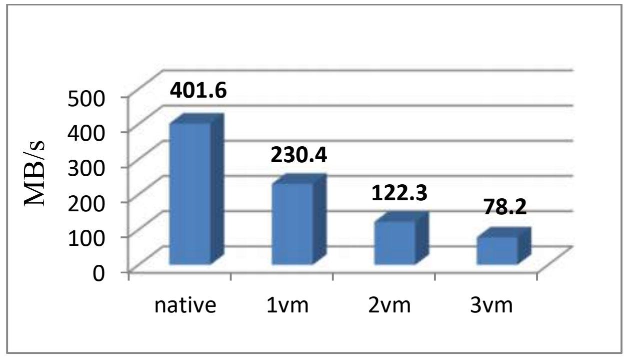


Fig. 4. Fileserver.f workload test results

The characteristics of the fileserver.f workload with fifty threads, are that both random and the sequential components dominate, but there is also a large number of I/O requests and much heavier data flow. A general notion is that, in the case for one virtual machine instance, the ESXi is significantly weaker than in the case of the native host OS. In the case of fileserver.f workload, when two virtual machines are instanced, the performance is further degraded by approximately the same amount as in the previous case (with one virtual machine). Instancing a third virtual machine brings very little performance degradation (Figure 4).

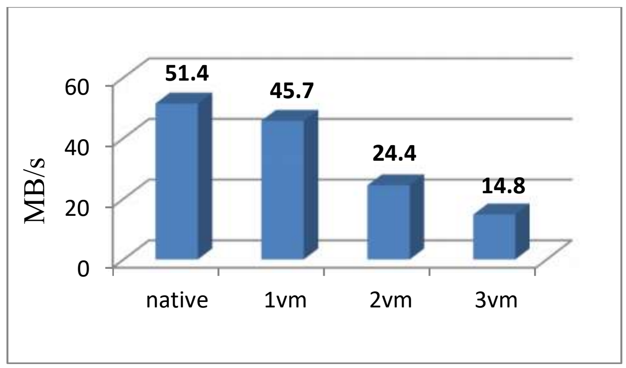


Fig. 5. Varmail.f workload test results

The characteristics of the varmail.f workload with our specification (16 threads) are that the components of random read and write are dominating, while the sequential components are not present, as it is shown in Figure 5. The special characteristic is that the components of the random write are synchronous, so each write will end up on the disk. The general notion for one instance is that performances of ESXi virtual machine are close to the native host OS. However, synchronous entries cancel the effects of cashing, so there are minor differences between native host OS and one instance of ESXi virtual machine. In the case of varmail.f workload, the instantiation of more than one ESXi virtual machines does not bring significant performance degradation.

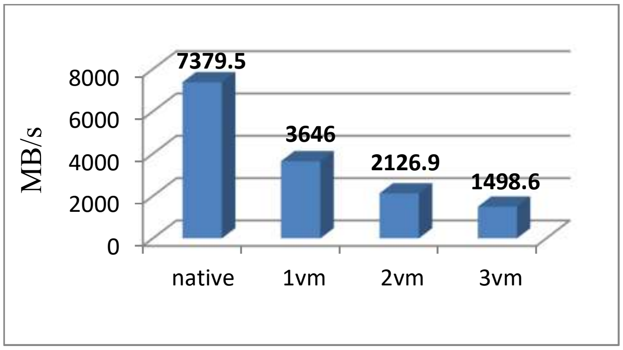


Fig. 6. Randomfileaccess.f workload test results

The characteristics of the randomfileaccess.f workload with five threads are that random reads dominate, while the sequential components are not present as shown on figure 6. We set up this workload so that cache would not be used. As we observe, the native host OS performs significantly better compared to one instance of ESXi virtual machine running. After starting second and third instances of ESXi virtual machines, we were able to observe that performance degradation is still present but it is not too significant.

The acquired benchmark results are fully expected and in line with the theoretical assumptions. The ESXi hypervisor and the hardware assisted full virtualization model show clear limitations on the data flow, in particular with the increase in the number of active virtual machines that cause even greater sharing of processors' resources and its increased use for hardware simulation. The addition of new instances of virtual machines is even more decreasing the achieved data flow, which means that new virtual machines cannot be added to the indefinite, as the performance of the whole system degrades per virtual machine added.

VIII. Conclusion

The introduction of virtualization has led to major changes in the use and deployment of information technology. Virtualization technology has a significant impact on reducing hardware investment as well as reducing operating costs, while also providing many additional benefits other than server consolidation. The great expansion of cloud computing in recent years has also contributed to the accelerated development of virtualization technologies and in the foreseeable future virtualization will always have an increased application in information technologies. It is also reasonable to expect, given the development of information technology today, that virtualization techniques will continue to improve and that the performance gap between virtualized systems and native systems will narrow in the future.

The results of our measurements showed that a native machine works convincingly better in most cases than a virtual machine based on ESXi hypervisor and full virtualization, as we assumed in our hypothesis of expected behaviour. Virtual machines running on the ESXi hypervisor have lower performance than a native machine, and in three of the four tests the performance degradation is approximately 50% when we have only one instance of a virtual machine running. The performance degradation is even greater with the introduction of more virtual machines. In one of the tests (mail server), the performance degradation between a native and a single virtual machine is not large, but with the introduction of new virtual machines into the test environment, the performance degradation of virtual machines becomes extremely pronounced. Future research may include a different approach where instead of comparing native machine vs. virtualized one, we compare different types of similarly structured virtualized machines or systems. With this research, we have proven that virtual systems still cannot reach the performance of non-virtualized systems, but as technologies evolve at an accelerated pace, we hope that in the future the performance of virtual machines will reach or be equal to regular non-virtualized machines.

Acknowledgement

The work presented in this paper has partially been funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

Literature

[1] C. Jiang, B. Luo, J. Wang, J. Zhang, Y. Wang, W. Shi, "Energy efficiency comparison of hypervisor," Proc. 2016 Seventh International Green and Sustainable Computing Conference *(IGSC)*, pp. 1-8, 2016.

[2] Correia, "Hypervisor based server virtualization" in *Encyclopaedia of information Science and Technology*, IGI Global, 2015, pp. 1182-1187.

[3] W. Huang, J. Liu, B. Abali, D. K. Panda, "A case for high performance computing with virtual machines," Proc. of the International Conference on Supercomputing. ACM, pp. 125-134, 2006.

[4] G. Casale, S. Kraft, D. Krishnamurthy, "A model of storage I/O performance interference in virtualized systems," Proc. of 31st International Conference on Distributed Computing Systems Workshops, pp. 34-39, 2011.

[5] J. Che, Q. He, Q. Gao, D. Huang, "Performance Measuring and Comparing of Virtual Machine Monitors," Proc. of 5th International Conf. Embedded and Ubiquitous Computing. EUC2008, Vol. 2, Piscataway, NJ, USA, pp. 381–386, 2008.

[6] A. Bhatia and G. Bhattal, "A comparative study of various hypervisors performance," International Journal of Scientific and Engineering Research, vol. 7, no. 12, pp. 65-71, 2016.

[7] J. Hwang, S. Zeng, F. Y. Wu, and T. Wood, "A component-based performance comparison of four hypervisors," Proc. of 2013 IFIP/IEEE International Symposium. IEEE, pp. 269-276, 2013.

[8] Pousa, Duarte; Rufino, José, "Evaluation of type-1 hypervisors on desktop-class virtualization hosts," IADIS International Journal on Computer Science and Information Systems. ISSN 1646-3692. 12:2, p. 86-101, 2017.

[9] H. Kazan, L. Perneel, M. Timmermann, "Benchmarking the performance of Microsoft Hyper-V server, VMWare ESXi and Xen hypervisors," Journal of Emerging Trends in Computing and Information Sciences, vol. 4, no. 12, pp. 922-933, 2013.

[10] Filebench project, Available: https://github.com/filebench/filebench/wiki

[11] VMware vSphere Documentation: https://docs.vmware.com/en/VMware-vSphere/index.html

[12] VMware ESXi 6.7 project: https://docs.vmware.com/en/VMware-vSphere/6.7/vsphere-esxi-vcenter-server-67-storage-guide.pdf.pdf

[13] VMware, Inc. white paper. "Virtualization Overview". https://www.vmware.com/pdf/virtualization_considerations.pdf

[14] VMware, Inc. white paper. "The Architecture of VMware ESXi," p. 3, 2020. https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/techpaper/ESXi_architecture.pdf

ESXi and Proxmox: FileSystem Performance Comparison for Type-1 Hypervisors

Borislav Đorđević, *Member, IEEE*, Valentina Timčenko, *Member, IEEE,* Nenad Nedeljković, Nikola Davidović

Abstract **— This paper presents the comparison of two representatives of type 1 hypervisors: Proxmox VE and VMware ESXi. Hypervisor acts like a lightweight operating system and runs directly on the host's hardware. The measurements are carried out on the same server and under the equivalent conditions, with the Linux Ubuntu 20.10 as the guest operating system using the Filebench 1.5-alpha1 software. The goal of this paper is to show an impact of different number of virtual machines on the performances of various file system and highlight the best combination. The results have been illustrated in graphical form.**

Keywords — Virtualization; Hypervisor; Proxmox; ESXi; Filebench; virtual machine

I. INTRODUCTION

The virtualization is considered as one of the most important topics in IT. It allows a single computer/server to use multiple operating systems simultaneously. It also helps in reducing the costs, because they can run multiple different services on a single server, leading to more efficient server utilization, easier system maintenance, and reduced hardware. As the power of a computer unit has significantly increased since 1960s when the IBM's presented its visionary idea of virtualization, this solution became popular in system implementation and maintenance [1].

There are several approaches for virtualization in IT environments: hardware, software, desktop, data, network, memory, storage, etc. The hardware virtualization implies the use of a hypervisor, which is an additional layer that lies between hardware and operating system (OS) and makes a slight delay for when accessing the resources for virtualized environment, providing lower performances when compared to bare metal or non-virtualized system [2], [3].

Actually, hypervisor is specialized firmware and/or software installed on single hardware that allows hosting of the VMs.

There are two types of hypervisors (Figure 1): type 1, that is executed directly on hardware and manages guest OSs (ESXi, Proxmox); and type 2 that is executed on the host OS (VirtualBox, VMware Workstation).

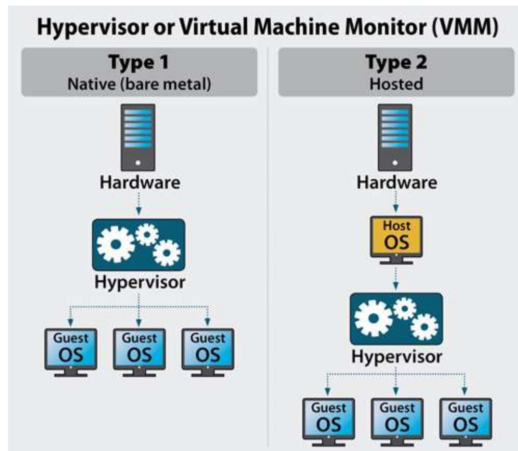


Figure 1. Hypervisor types and differences [4]

As the type 1 hypervisor has direct access to hardware, while type 2 hypervisor accesses hardware through host OS, we assume that type 1 hypervisor provides more scalability, reliability, and better performance [5].

II. RELATED WORK, OBJECTIVE AND MOTIVATION

This research is focused on the performance comparison of two type-1 hypervisors and results analysis. Since virtualization is the primary solution for systems ranging from small firms to large corporations, the arising question is: what is the best solution on the market? Some recent research addresses this issue from different perspectives, mostly considering VMware, KVM and Hyper-V hypervisors, and basing the results on Filebench or Bonnie++ [6]. This paper can provide a new picture of the situation since almost no research has focused on the Proxmox solution versus a commercial solution such as ESXi.

The primary goal of this paper is to compare performance using ESXi and Proxmox hypervisors on identical hardware, same VM parameters and the same guest OS – Linux Ubuntu 20.10 with ext4 as main file system (FS). Also, the disk we are testing has contained one of the three FSs: ext4, xfs or btrfs. Since we have used a Filebench workloads for testing, our idea was to find the best FS for each test. Selected workloads are: varmail, webserver and fileserver.

We have defined the mathematical model, measured the performances and interpreted the obtained results based on the mathematical model and hypotheses.

Borislav Đorđević - Institute Mihajlo Pupin, Volgina 15, 11000 Belgrade, Serbia, (borislav.djordjevic@pupin.rs)
Valentina Timčenko - Institute Mihajlo Pupin, Volgina 15, 11000 Belgrade, Serbia (valentina.timcenko@pupin.rs)
Nenad Nedeljković - VISER, School of Electrical and Computer Enginering of Applied Studies, Belgrade, Serbia (nedeljkovic1nenad@gmail.com)
Nikola Davidović – University of East Sarajevo, Faculty of Electrical Engineering, Vuka Karadzica 30, 71123 East Sarajevo, RS, BiH, (nikola.davidovic@etf.unssa.rs.ba)

III. MATHEMATICAL MODEL

Variable T_W is calculated in accordance with the equation (1), and shows the total processing time for each workload.

$$T_W = T_{RR} + T_{SR} + T_{RW} + T_{SW} \quad (1)$$

Variables T_{RR} and T_{SR} represent random and sequential read time, and T_{RW} and T_{SW} random and sequential write time. There is an expected access time for every specific workload for the FS, which include following components:

$$T_{WORKLOAD} = T_{DIR} + T_{META} + T_{FL} + T_J + T_{HK} \quad (2)$$

$T_{WORKLOAD}$ represents the overall time for finishing all operations on the current workload, T_{DIR} the time needed to run all directory-related operations, T_{META} the time needed to complete all metadata operations, T_{FL} the time needed to go through all free lists operations, T_{FB} the time needed to carry out direct file blocks operations, T_J the time needed to complete journaling operations and T_{HK} the time needed to run housekeeping operation within the FS [7].

We have two candidates whose performances we compare:

1. Proxmox + VMs (guest OS).
2. ESXi + VMs (guest OS).

1. Proxmox + VMs (guest OS): The time to process the generated workload ($T_{W(Proxmox)}$) in this case depends on the benchmark interaction with guestOS FS, the characteristic of the FS, Virtual Hardware processing and the virtualization processing component of the Proxmox hypervisor (PVE-proc) is calculated in accordance to the following formula:

$$T_{W(Proxmox)} = f(BENCH,\ guestOS\text{-}FS,\ VH\text{-}proc,\ PVE\text{-}proc,\ hostOS\text{-}FS) \quad (3)$$

2. ESXi + VMs (guest OS): The time to process the generated workload ($T_{W(ESXi)}$) in this case depends on the benchmark interaction with guestOS FS, the characteristic of the FS, Virtual Hardware processing and the virtualization processing component of the ESXi hypervisor (ESXi-proc) is calculated in accordance with the following formula:

$$T_{W(ESXi)} = f(BENCH,\ guestOS\text{-}FS,\ VH\text{-}proc,\ ESXi\text{-}proc,\ hostOS\text{-}FS) \quad (4)$$

Since we are using the same settings for VMs on both hypervisors, the virtualization processing component will depend on the virtualization type and hypervisor processing as provided in the following formula:

$$PVE\text{-}proc = f\,(virt\_type,\ hyp\_proc) \quad (5)$$
$$ESXi\text{-}proc = f\,(virt\_type,\ hyp\_proc) \quad (6)$$

We are predicting the following:

- Based on the practical experience, it is expected that ESXi will produce better performance.
- Multiple VMs to have a significant performance drop compared to just one VM.

IV. FILE SYSTEMS

Linux FS is generally a built-in layer of a Linux OS used to handle the data management of the storage.

A. EXT4

The ext4 (fourth **ext**ended filesystem) is a journaling FS for Linux, and is developed as the extension of the ext3 [8]. It has the following characteristics [9]:

- Maximum FS size of up to 1 EB and maximum file size of nearly 16 TB.
- Hashed B-tree organizes and finds directory entries.
- Online defragmentation tool (e4defrag), which performs defragmentation of individual files or the whole FS.
- Easily detectable corruptions of files by metadata checksumming.

B. XFS

XFS is a high-performance journaling FS created by Silicon Graphics, Inc (SGI) in the last decade of 20th century [10]. It has the following characteristics [9]:

- Maximum FS size and maximum file size of nearly 8 EB.
- B+ tree organizes and finds directory entries.
- Delayed allocation for minimizing fragmentation and increasing performance.
- Implemented direct I/O for high throughput and non-cached I/O for DMA devices.

C. BTRFS

Btrfs ("better FS", "b-tree F S") is a copy-on-write (COW) FS based on B-trees. It was initially designed at Oracle Corporation in 2007 for the use in Linux [11]. It has the following characteristics [9]:

- Maximum FS size and maximum file size of nearly 16 EB.
- B-tree organizes and finds directory entries.
- Online defragmentation, offline FS check.
- Background based fixing errors on redundant files.

V. VMWARE ESXI AND PROXMOX

ESXi is an enterprise-class, type-1 hypervisor developed by VMware for deploying and serving VMs (Figure 2). It runs directly on hardware and significantly improves system performance [12]. The major part of architecture is VMkernel and processes that run on top of it. VMkernel has control of all hardware devices on server, manages resources and handles system processes. It receives requests from VMs for resources and presents the requests to the physical hardware [13].

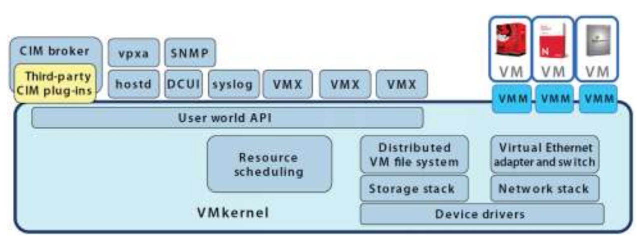


Figure 2. ESXi architecture [14]

The main processes that run on top of VMkernel are: [13]

• Direct Console User Interface (DCUI) — the low-level configuration and management interface, accessible through the console of the server, used primarily for initial basic configuration.

• The VM monitor, which is the process that provides the execution environment for a VM, as well as a helper process known as VMX. Each running VM has its own VMM and VMX process.

• Various agents used to enable high-level VMware Infrastructure management from remote applications.

• The Common Information Model (CIM) system: CIM is the interface that enables hardware-level management from remote applications via a set of standard APIs.

VMware uses VMFS. It is a special high-performance clustered FS. The main feature of this segment is ability to be shared by being simultaneously mounted on multiple servers. The VMFS datastore can be extended to span over several physical storage devices that include SAN LUNs and local storage. This feature allows you to pool storage and gives you flexibility in creating the datastore necessary for your virtual machines. [12]

Proxmox **V**irtual **E**nvironment – PVE (Figure 3) is a bare-metal hypervisor (runs directly on the hardware), to run VMs and containers. It is an open-source project, developed and maintained by Proxmox Server Solutions GmbH. For maximum flexibility, they implemented two virtualization technologies: full virtualization with KVM (Kernel-based Virtual Machine) and container-based virtualization (LXC) [15].

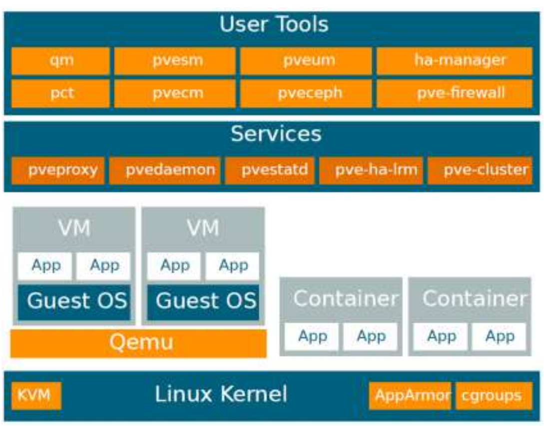


Figure 3. Proxmox architecture [15]

Proxmox uses a Linux kernel and is based on the Debian GNU/Linux Distribution. The source code is released under the GNU Affero General Public License, version 3. KVM was the first hypervisor to become part of the native Linux kernel (2.6.20). It is implemented as a kernel module, allowing Linux to become a hypervisor simply by loading a module. Benefits from the changes to the mainline version of Linux is optimization of hypervisor and the Linux guest Oss [16].

Proxmox natively supports running LXC (LinuX Containers) containers from the UI. These are similar to docker containers but behave more like a traditional VM.

Performance of KVM virtualization was the focus of this paper.

The main features for Proxmox VE [17]:

- Live migration;
- High availability;
- Scheduled backup;
- Command-line (CLI) tool;
- Flexible storage;
- OS template.

VI. TESTING

The assumption of adequate testing is the application of a single hardware configuration, the same OS, and measurement methodology for all tests. The used server configuration has respectable hardware components although it is does not represent the latest technology.

The OS used is Ubuntu version 20.10, the latest instalment of Linux distribution (Table 1). During the installation process, we opted for minimal installation option which installs only essential packages and programs. The system disk uses EXT4 while the test disk is EXT4, XFS, or BTRFS.

All tests were performed using Filebench tool. Latest release of Filebench software was installed following instructions provided on the official GitHub repository of this project. Filebench is a program designed to measure the performance of FS and storage, and it can generate multiple workload types that simulate environments when using certain servers/services such as mail, web, file, database, etc. [18]. Before starting any tests, we made sure that all available updates were installed. Each VM was given 4 GB of RAM and 4 CPU cores.

TABLE I
SERVER TEST ENVIRONMENT

HP ProLiant DL380 G7	
Component	***Characteristic***
CPU	2 x Intel Xeon E5540 QuadCore 2.53GHz
RAM	32GB DDR3
Storage Controllers	HP Smart Array P410i
Hard Drive 1	HP 10K SAS 146GB(DG0146)
Hard Drive 2	HP 7.2K SAS 500GB(MM0500)
PVE hostOS-FS	ext4
ESXi hostOS-FS	VMFS

The VM parameters are shown in Table 2. All used VMs have identical characteristics.

TABLE II
VIRTUAL MACHINE PARAMETERS

Component	***Characteristic***

vCPU	4
RAM	4GB
Disk	12GB + 32GB
OS	Linux Ubuntu 20.10
FS	ext4
Tested FS	ext4/xfs/btrfs

The focus of this paper is on measuring disk performance by comparing two hypervisors combined with three different FSs using 1, 2, or 3 VMs at the same time. It is expected that, as the number of VMs increases, performance will decline significantly in any combination.

Filebench is a very powerful and very flexible tool able to generate a variety of FS - and storage-based workloads. It implements a set of basic primitives like *create file*, *read file*, *mkdir*, *fsync* and uses WLM (the Workload Model Language - WML) to combine these primitives in complex workloads [18].

The files used for our benchmark were *varmail.f, webserver.f,* and *fileserver.f.* Those files are included in the Filebench software installation package, and were minimally edited to suit our needs.

The duration of each the tests was set to 120 seconds, which is the only change we made in *.f files with the goal of making the most realistic results. During the test execution, it was ensured that the impact of any external subject on system components was reduced to the minimum. The benchmark is run 3 times and the average value of the test is taken as final.

First, Proxmox VE was installed on server and nine VMs were generated, 3 for every FS. Tests were conducted in a way that one VM was first started and measured, then 2 and 3 VMs simultaneously. After that, disk is formatted and ESXi was installed. By the same principle, everything is applied to ESXi. From the generated data, the final conclusions were made by calculating the average values of the results.

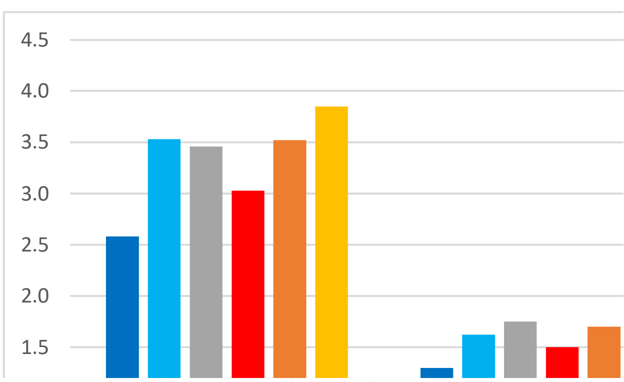


Figure 5. Varmail workload test results

TABLE III
BENCHMARK VARMAIL RESULTS

Varmail	1VM - (MB/s)	2VM - (MB/s)	3VM - (MB/s)
esxi - ext4	2.6	1.3	0.9
esxi - xfs	3.5	1.6	1.2
esxi - btrfs	3.5	1.8	1.0
pve - ext4	3.0	1.5	1.1
pve - xfs	3.5	1.7	1.2
pve - btrfs	3.9	1.9	1.5

Figure 5 and Table 3 show Varmail test results. Varmail emulates I/O activity of a simple mail server that stores each e-mail in a separate file (/var/mail/ server). The workload consists of a multi-threaded set of create-append-sync, read-append-sync, read and delete operations in a single directory. 16 threads are used by default [19].

For the Varmail workload, which is characterized by the dominant random reads and random writes, where random writes are represented by the synchronous transfers covered by equations (3) and (4), the main differences are components 3 (VH-proc), 4 (hypervisor-proc) and 5 (hostOS-FS).

When looking at the number of VMs, the combination of pve-btrfs was the best in each category, while esxi-btrfs and esxi-xfs had the same overall results with the ESXi hypervisor. We can conclude that the BTRFS FS is the best choice for a mail server.

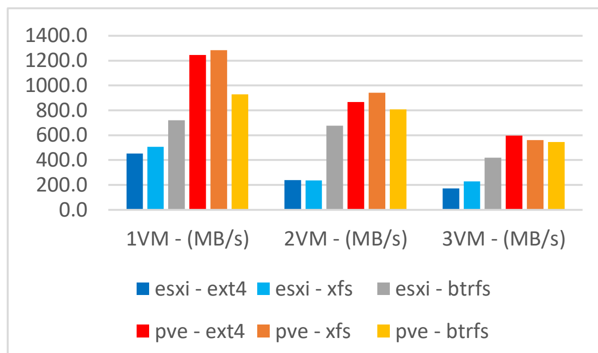


Figure 6. Webserver workload test results

TABLE IV
BENCHMARK WEBSERVER RESULTS

Webserver	1VM - (MB/s)	2VM - (MB/s)	3VM - (MB/s)
esxi - ext4	453.5	238.4	171.2
esxi - xfs	507.2	235.1	228.3
esxi - btrfs	720.0	677.5	419.1
pve - ext4	1243.9	864.9	595.0
pve - xfs	1284.1	940.5	561.8
pve - btrfs	928.8	808.0	544.7

Figure 6 and Table 4 show Webserver test results. Webserver emulates simple web-server I/O activity and produces a sequence of open-read-close on multiple files in a directory tree plus a log file append. 100 threads are used by default [19]. The Webserver workload is characterized by a dominant random read component as covered in equations (3) and (4), while the main differences are components 3 (VH-proc) and 5 (hostOS-FS). In both cases, VH-proc is Full-Hardware virtualization, but in Proxmox it is realized through QEMU. A large difference in performance in favor of Proxmox was observed in this test. The overall results of pve-xfs is 2.87 times better than esxi-xfs.

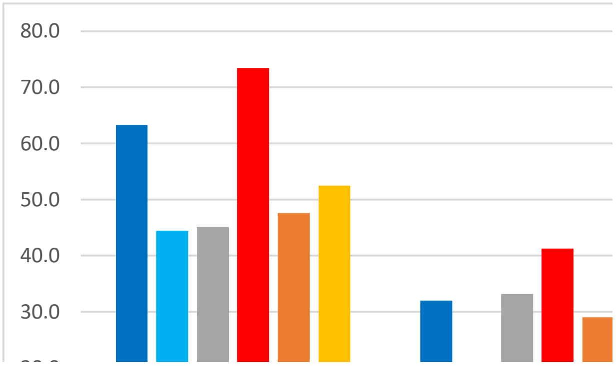


Figure 7. Fileserver workload test results

TABLE V
BENCHMARK FILESERVER RESULTS

Fileserver	1VM - (MB/s)	2VM - (MB/s)	3VM - (MB/s)
esxi - ext4	63.3	32.0	17.4
esxi - xfs	44.4	16.7	13.8
esxi - btrfs	45.1	33.2	10.7
pve - ext4	73.4	41.3	21.4
pve - xfs	47.6	29.0	18.3
pve - btrfs	52.5	26.4	20.4

Figure 7 and Table 5 show Fileserver test results. Fileserver - Emulates simple file-server I/O activity. This workload performs a sequence of creates, deletes, appends, reads, writes and attribute operations on a directory tree. 50 threads are used by default [19].

For the Fileserver workload, which is characterized by all kinds of data transfers, when considering equations (3) and (4), the main difference is component 5 (hostOS-FS). As ESXi uses VMFS, which is a clustered FS and represents a higher level of abstraction, while Proxmox uses EXT4, and the best FS in this test was EXT4, we conclude that this ruled in favor of Proxmox.

As in the previous two tests, this time too Proxmox came out as the winner but with a slightly smaller difference. We also have a match in the choice of FS: EXT4 gave the best overall results in both hypervisors.

VII. CONCLUSION

In this paper, we tested two respectable type 1 hypervisors: the commercial VMware ESXi solution and the open-source solution - Proxmox. Although it was expected that, due to its importance and big impact in the IT world, ESXi would provide better results, this did not happen. Proxmox won each comparator hypervisor + file system test. This was best seen during the webserver test where they were better almost 3 times and the third (VH-proc) and fifth (hostOS-FS) components of formulas (3) and (4) came to the fore.

If we only look at the performance of the FS, we get an interesting distribution. EXT4 performed best on fileserver, XFS on webserver, and BTRFS on varmail test.

For all 3 workloads we noticed that Proxmox is significantly better than ESXi. In the context of formulas (3), (4), (5), (6), we consider that the first two components, BENCH and guestOS-FS, in equations (3) and (4) have the same effect on for both hypervisors. The 3rd and 4th components, VH-proc, PVE-proc and ESXi-proc, differ significantly, where we notice that Proxmox is better. However, the main reason for Proxmox's victory is the 5th component (hostOS-FS). ESXi used a higher level of abstraction such as VMFS which slowed it down in this case, while Proxmox used a basic level of FS such as EXT4.

When we summarize all the test results, the used virtual machine operating system and hypervisors hostOS-FS, we can say that Proxmox is more optimized for Linux distribution.

The Proxmox virtualization system can be particulary useful for people starting their own business in small steps, without requiring additional costs. This does not mean that large companies do not use it. As already mentioned, this is an Open-Source solution and help for some of the possible problems can be found in a community where the number is unknown. If you still want to be insured, you can subscribe to the team of people behind this solution - Proxmox Server Solutions GmbH on more than favorable terms.

Interesting ideas for future work and research is to add fast Solid State Disks, comparative analysis of hypervisors using container virtualization or testing a different hypervisor such as Xen and Microsoft Hyper-V to determine which one achieves the best results.

ACKNOWLEDGMENT

The work presented in this paper has partially been funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

REFERENCES

[1] S. Meier, *IBM Systems Virtualization: Servers, Storage, and Software*, First edition, Redpaper, 2008.

[2] F. Bazargan, C. Yeun, J. Zemerly, "State-of-the-Art of Virtualization, its Security Threats and Deployment Models", International Journal for Information Security Research. 3. 10.20533/ijisr.2042.4639.2013.0039, 2013.

[3] N. Yaqub, "Comparison of Virtualization Performance: VMware and KVM", Master Thesis, Department of Informatics, Uviversity of Oslo, Norway, 2012.

[4] C. Taylor, 2020, What is a Hypervisor Server?, accessed 14 May 2021, https://www.serverwatch.com/virtualization/hypervisor-server/

[5] P. Vasconcelos, F. Araújo, G. Freitas, T. Marques, "KVM, OpenVZ and Linux Containers: Performance Comparison of Virtualization for Web Conferencing Systems", International Journal of Multimedia and Image Processing. 6. 10.20533/ijmip.2042.4647.2016.0039, 2016.

[6] B. Đorđević, N. Maček, V. Timčenko, "Performance Issues in Cloud Computing: KVM Hypervisor's Cache Modes Evaluation", Vol. 12, No. 4, pp 147-165, 2015. http://uni-obuda.hu/journal/Dordevic_Macek_Timcenko_60.pdf,

[7] D. Vojnak, B. Đorđević, V. Timčenko, S. Štrbac, "Performance Comparison of the type-2 hypervisor VirtualBox and VMWare

Workstation”, 27th Telecommunications Forum (TELFOR), Belgrade, Serbia, pp. 1-4, doi: 10.1109/TELFOR48224.2019.8971213, 2019.

[8] THE EXT4 FILE SYSTEM, accessed 16 May 2021, https://access.redhat.com/documentation/en-us/red_hat_enterprise_linux/7/html/storage_administration_guide/ch-ext4

[9] D. Vujičić, D. Marković, B. Đorđević, S. Ranđić, “Benchmarking Performance of EXT4, XFS and BTRFS as Guest File Systems Under Linux Environment”, 2016.

[10] THE XFS FILE SYSTEM, accessed 16 May 2021, https://access.redhat.com/documentation/en-us/red_hat_enterprise_linux/7/html/storage_administration_guide/ch-xfsBtrfs

[11] O. Rodeh, J. Bacik, C. Mason, (2013). “BTRFS: The linux B-tree filesystem”, ACM Transactions on Storage (TOS). 9.10.1145/2501620.2501623, 2013.

[12] E. Sosa, *Mastering VMware NSX® for vSphere®*, First edition, Indianapolis, Indiana, John Wiley & Sons, 2020.

[13] VMware, “The Architecture of VMware ESXi”, Technical White Paper, 2007.

[14] B. Laurent, 2020, VMs Everywhere, accessed 24 May 2021, https://littlecorner.info/post/virtualization/

[15] Proxmox Server Solutions Gmbh, “Proxmox VE Administration Guide”, 2021.

[16] S. Aiiy, “Comparative analysis of proxmox VE and xenserver as type 1 open source based hypervisors”, International Journal of Scientific and Technology Research. 7. 72-77, 2018.

[17] S. Cheng, *Proxmox High Availability*, First edition, Birmingham, England, Packt Publishing, 2014.

[18] V. Tarasov, E. Zadok, S. Shepler, “Filebench: A Flexibile Framework for File System Benchmarking”, Vol. 41, No. 1, 2016.

[19] Filebench, 2017, accessed 23 April 2021, https://github.com/filebench/filebench/wiki

Snort IDS system visualization interface

Nadja Gavrilovic, Vladimir Ciric, Nikola Lozo
University of Nis, Faculty of Electronic Engineering, Nis, Serbia

***Abstract*—Over the past decades, the rapid Internet development and the growth in the number of its users have raised various security issues. Despite numerous available security tools, the exchange of data over the Internet is becoming increasingly insecure. For this reason, it is of great importance to ensure the security of the network in order to enable the safe exchange of confidential data, as well as their integrity. One of the most important components of network attack detection is an Intrusion Detection System (IDS). Snort IDS is a widely used intrusion detection system, which logs alerts after detecting potentially dangerous network packets. The next step in successful network protection is the analysis of logged alerts in search of deviations from normal traffic that may indicate an intrusion. The goal of this paper is to design and implement a visualization interface that graphically presents alerts generated by Snort IDS, classifies them according to the most important attack parameters, and allows the users to easily detect possible traffic irregularities. An environment in which the system has been tested in real-time is described, and the results of attack detection and classification are given. One of the detected attacks is analyzed in detail, as well as the method of its detection and its possible consequences.**

***Index Terms*—IDS, snort, network intrusion detection, visualization interface**

Nadja Gavrilovic is with the Faculty of Electronic Engineering, University of Nis, Aleksandra Medvedeva 14, Nis, Serbia (e-mail:nadja.gavrilovic@elfak.ni.ac.rs).

Vladimir Ciric is with the Faculty of Electronic Engineering, University of Nis, Aleksandra Medvedeva 14, Nis, Serbia (e-mail:vladimir.ciric@elfak.ni.ac.rs).

Nikola Lozo is with the Faculty of Electronic Engineering, University of Nis, Aleksandra Medvedeva 14, Nis, Serbia (e-mail:nikolalozo@elfak.rs).

Performance comparison of homomorphic encryption scheme implementations

Goran Đorđević, *AET Europe The Netherlands, ETF Beograd, Milan Marković, Panevropski Univerzitet Apeiron Banja Luka, Pavle V. Vuletić, ETF Beograd*

Abstract **— Homomorphic Encryption allows third party to receive encrypted data and perform arbitrarily computations on that data while it remains encrypted, despite not having the secret decryption key. This enables many new secure applications in cloud environments. For a long time, a key issue with the homomorphic encryption was its low performance which made it unusable in production environments. Advances in the last ten years in the field of homomorphic encryption resulted in several new schemes and software libraries which implement them. These homomorphic schemes have improved performance, but there is still a question whether the improvements would justify their use in production environments. In this paper we evaluated features and performances of several new homomorphic encryption mechanisms: BGV, BFV and CKKS.**

Keywords **— Homomorphic Encryption; Performance; Secure Multiparty Computation.**

I. INTRODUCTION

Homomorphic encryption allows computations on ciphertext without the knowledge of the secret key, or more precisely it allows performing computations on the encrypted data, without decrypting them [1]. Homomorphic encryption allows a third party (e.g., cloud, service provider) to perform certain computable functions on the encrypted data while preserving the features of the function and format of the encrypted data and without being able to see its content. Homomorphism of the first asymmetric encryption algorithms (RSA) over some mathematical operations (e.g. multiplication) was known since these algorithms were invented almost fifty years ago. Such schemes which support partial set of mathematical operations are known as partially homomorphic. Cryptographic mechanisms that support arbitrary level of computations on ciphertext (multiplication, addition, rotation) without the knowledge of the secret keys are known as Fully Homomorphic Encryption (FHE) systems. The increased popularity of cloud-based services on one side and the need to preserve data privacy led to the new interest in homomorphic encryption research which would enable secure multiparty computation in the cloud environment. One could imagine the use of AI or machine learning algorithms on the data which is encrypted and invisible to the AI system provider, thus preserving data privacy only for the data owner. An example of such a scenario where homomorphic encryption mechanisms are deployed is given in Figure 1. In this example the user sends and stores the data in the encrypted form on the cloud server. The data is processed on the server in the encrypted form, and the results which remain in the encrypted form are sent back to the user who can decrypt the data and use the result.

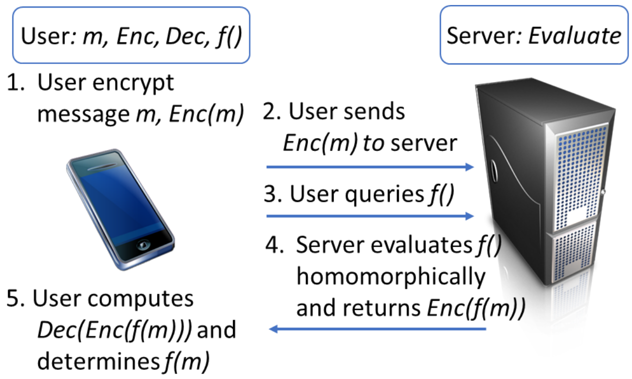


Fig. 1. An example of client-server HE scenario

The biggest obstacle for the use of homomorphic encryption schemes was the fact that there were no FHE mechanisms which had reasonable performance. Computations were by several orders of magnitude slower than the operation on unencrypted data which made any such solution too resource expensive. However, in the last ten years a breakthrough happened in the area and a set of new homomorphic encryption schemes emerged. Modern fully homomorphic encryption schemes use complex algorithms on lattice structures and Ring-LWE (Ring Learning With Errors) mechanism [2]. In addition to the homomorphic property, it is believed that these algorithms are resistant to quantum computer attacks because nowadays there are no known algorithms that would use the properties of quantum computers to break these algorithms in polynomial time. Following the appearance of new fully homomorphic encryption schemes, a set of programming APIs and libraries which implement different schemes emerged as well. In this paper we are assessing the set of capabilities and performance of three homomorphic encryption schemes (BGV, BFV, CKKS) and are discussing the suitability and constraints of these schemes for use in the cloud-based environments for secure multiparty computations. Performance assessment of the new FHE schemes has not been explored a lot in the literature. We believe that this paper will provide a better insight into the current state of the work on FHE and its suitability for real use case scenario deployments.

The paper is organized as follows. Section II gives an overview of the related work in the field of the performance evaluation of the FHE schemes. The most important properties of homomorphic encryption and the classification of the homomorphic encryption schemes are presented in Section III. The main features and description of modern HE schemes: BGV [3], BVF [4] and CKKS [5] are elaborated in Section IV. In Section V is shown results of experimental analysis. Conclusions are given in Section VI.

II. RELATED WORK

Related work about the FHE schemes is spread across the papers in the relevant sections, while this section contains only those papers which were dedicated to FHE performance evaluation. Experimental results related to BGV scheme with value of ciphertext modulus q=130 are given in [1]. Viand et al. in [6] compare the features of Palisade, Microsoft SEAL, and HELib homomorphic encryption libraries. In addition, this paper gives statistical compiler tests of BVF scheme implemented in the SEAL library in a graphical form without presenting precise numerical values. Melchor et al. [7] compared the performance of three libraries HELib, SEAL and FV-NFLlib for large plaintext moduli of up to 2048 bits. Finally, Lepoint et al. [8] compare the performance of two older homomorphic schemes. Unlike the previous work, in this paper we give experimental results for BGV with broader range of values ciphertext modulus q and results for other modern homomorphic schemes: BFV and CKKS that are not covered in [1].

III. PROPERTIES OF HOMOMORPHIC ENCRYPTION

There are four main types of homomorphic schemes [1]:

- Partially Homomorphic Encryption (PHE). The PHE scheme enables either any number of addition or any number of multiplication operations over encrypted data.
- Somewhat Homomorphic Encryption (SHE) allows both addition and multiplication, but it can perform a limited number of operations. “Somewhat” means it works for some functions *f*.
- Fully Homomorphic Encryption. The scheme allows any number of addition or multiplication operations. “Fully” means it works for all functions *f*. An FHE scheme can evaluate unbounded depth.
- Levelled Homomorphic Encryptions (LHE). This scheme can evaluate arbitrary polynomial-size circuits.

Homomorphic Encryption should support two main homomorphic operations:

- Additive Homomorphic Encryption;
- Multiplicative Homomorphic Encryption.

Homomorphic encryption is additive, if [9]:

$$\text{Enc}\,(m_1 + m_2) = \text{Enc}\,(m_1) + \text{Enc}\,(m_2);\ \forall m_1, m_2 \in M.$$

Homomorphic encryption is multiplicative, if [9]:

$$\text{Enc}\,(m_1 * m_2) = \text{Enc}\,(m_1) * \text{Enc}\,(m_2);\ \forall m_1, m_2 \in M.$$

The most popular classes of homomorphic schemes are (given with their main properties):

- Boolean circuit (Fastest Homomorphic Encryption in the West (FHEW) [10] and Fast Fully Homomorphic Encryption over the Torus (TFHE) [11]):
  - Plaintext data are coded as bits;
  - Computations are performed by using Boolean circuits.
- Modular integer arithmetic (BGV, BFV):
  - Plaintext data are coded as integer modulo a plaintext;
  - Computations are expressed as integer modulo arithmetic.
- Approximate number arithmetic (CKKS):
  - Plaintext data are coded as real (or complex) numbers;
  - Computations are performed in a way similar to floating-point arithmetic but dealing with fixed-point numbers.

Modern HE mechanisms are based on usage of lattice cryptography with errors LWE [12]. Lattices have an important role in modern cryptography, especially in the context of the research on post-quantum cryptography. It is known that the factoring problem which was discovered to be solvable in polynomial time on a quantum computer by Shor can be applied to the widely used asymmetric cryptographic schemes (RSA, DH). At the moment of writing this paper there was no report in the literature which claimed that it can break lattice-based cryptographic algorithms using quantum computer algorithms.

The newest HE algorithms are applied structured lattices i.e. Ring-LWE mechanism [2]. The Ring-LWE reduces key length and computation time. The ring implementation is based on power-of-two cyclotomic rings:

$$\mathrm{R}_q = \mathbb{Z}_q \,/\, \langle \mathrm{x}^n + 1\rangle$$

The optimized Residue Number System (RNS) variants of algorithms show significant performance gain compared to their earlier respective implementations [13]. The RNS works with native (machine-word size) integers because it is faster than multi-precision integer arithmetic. It breaks rings of large bit-width integers into a parallel set of rings (<64-bit residues) allowing very efficient computation on 64-bit CPU architecture.

Large modulus q is represented as product of integers:

$$q = \prod_{i=1}^{k} qi$$

Modulus q is a functional parameter that determines how many computations are allowed without the appliance bootstrapping procedure [14].

One of the properties of the homomorphic encryption schemes is that they add noise to a ciphertext in the encryption process. Homomorphic operations (especially multiplication) increase the noise. If the noise becomes too large, the resultant ciphertext can become undecryptable. Noise budget is the total amount of noise that can be added until the decryption fails [15]. The bootstrapping is the procedure of "refreshing" a ciphertext by running the decryption function on it homomorphically, resulting in a reduced noise.

All considered homomorphic encryption schemes support the following homomorphic operations:

- Addition;
- Multiplication;
- Rotation.

IV. Homomorphic Schemes

The BGV scheme was proposed [3]. BGV is a levelled HE scheme, meaning that the parameters of the scheme depend on the multiplicative depth that the scheme is capable to evaluate. Multiplicative depth determines how many sequential multiplications can be performed.

The BFV scheme [4] is a homomorphic cryptographic scheme based on the Ring-LWE problem in a lattice.

The CKKS scheme [5] is known as Homomorphic Encryption for Arithmetic of Approximate Numbers (HEAAN). Supported operations in the scheme are shown in Figure 2. The CKKS scheme enables computations on vectors of complex values.

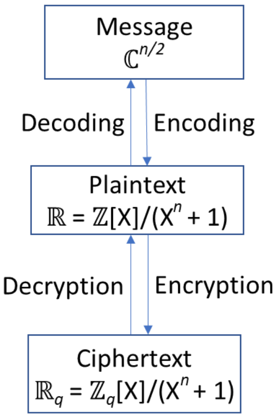


Fig. 2. Operations in CKKS

The CKKS is an approximate homomorphic encryption scheme with the following features:

- Dec (Enc(m)) ≈ m;
- Dec (ct_1 * ct_2) ≈ Dec (ct_1) * Dec (ct_2);
- Noise bounds are determined by the parameter set.

In the CKKS scheme noise is considered as a part of numerical error in approximate computation. It supports homomorphic rounding-off.

In all above-mentioned schemes the following homomorphic operations are implemented [16]:

- Public key encryption:
  $$PubEncrypt(pk, M) \rightarrow C$$
  The public encryption algorithm takes as input the public key (pk) of the scheme and any message M from the message space. The algorithm outputs a ciphertext C.
- Decryption:
  $$Decrypt(sk, C) \rightarrow M$$
  The decryption algorithm takes as input the secret key of the scheme (sk), and a ciphertext C. It outputs a message M from the message space.
- Homomorphic addition:
  $$EvalAdd(Params, ek, C_1, C_2) \rightarrow C_3$$
  EvalAdd is an algorithm that takes as input the system parameters *Params*, the evaluation key (ek), two ciphertexts C_1 and C_2, and outputs a ciphertext C_3.
- Homomorphic multiplication:
  $$EvalMult(Params, ek, C_1, C_2) \rightarrow C_3$$
  EvalMult is an algorithm that takes as input the system parameters *Params*, the evaluation key ek, two ciphertexts C_1 and C_2, and outputs a ciphertext C_3.

The evaluation key is needed to perform homomorphic operations over the ciphertexts. The evaluation key is used in in the following homomorphic operations: relinearization (multiplication) and rotation. Any entity that has only the evaluation key cannot learn anything about the messages from the ciphertexts only [16].

An example of homomorphic encryption with asymmetric key cryptography by using BGV [3], BVF [4], and CKKS [5] schemes is shown in Figure 3.

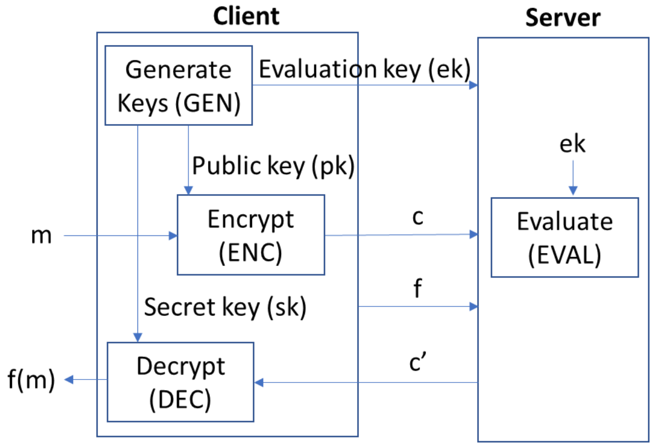


Fig. 3. Homomorphic encryption with asymmetric keys

V. Experimental Analysis

In the experimental analysis we evaluated the time needed for execution of the following homomorphic operations: Public key encryption (Table II), Decryption (Table III), Homomorphic addition (Figure 4), and Homomorphic multiplication (Figure 5). Homomorphic encryption libraries implement the above-mentioned cryptographic operations of a scheme and expose a higher-level API. We evaluated the use of the following homomorphic schemes:

- BGV,
- BVF and
- CKKS;

that are implemented in the following open-source libraries respectively:

- Microsoft SEAL [17];
- Palisade [14];
- HELib [18] [19].

HELib is a C++ open source library that implements both the BGV [3] and CKKS [5] homomorphic encryption schemes. HELib library, published in 2013 by Halevi and Shoup, was the first homomorphic encryption library.

Palisade [14] is multi-threaded library written in C++ 11. It uses the NTL library [20] to accelerate underlying mathematical operations. Palisade supports more schemes, including BFV, BGV, CKKS. It also supports multi-party extensions of certain schemes and other cryptographic primitives like Proxy Re-Encryption (PRE) and digital signatures [6].

Microsoft Simple Encrypted Arithmetic Library (SEAL) [17] is a homomorphic encryption library that allows additions and multiplications to be performed on encrypted integers or real numbers. Microsoft SEAL is written in C++11 and contains a .NET wrapper library for the public API. The latest available version 3.6.2 is developed in C++17.

Table I gives an overview of the publicly available open-source libraries with implemented HE algorithms. Palisade implements Boolean circuits Fully Homomorphic Encryption (FHE) schemes: FHEW and TFHE. In the FHE mechanisms it uses bootstrapping procedure [14] (noise refreshing procedure) with the application of the appropriate bootstrapping keys. The FHEW and TFHE schemes are not implemented in the HELib and Microsoft SEAL libraries.

TABLE I
HE ALGORITHMS IN OPEN-SOURCE LIBRARIES

Library/ HE scheme	Palisade	HELib	SEAL
BGV	√	√	
BFV	√		√
CKKS	√	√	√
FHEW	√		
Threshold FHE	√		

The homomorphic encryption code was executed on a PC with:

- 2194.84 MHz 8-core CPU;
- 16 GB RAM;
- Ubuntu 20.04 LTS.

Tables II and III and Figures 4 and 5 show the results of encryption, decryption, HE addition and HE multiplication tests respectively, where:

- Times in the last three columns (HE Library) are expressed in microsecond (μs);
- Each operation was executed 1000 times and the times presented are the times to execute 1000 iterations;
- We used 128-bit homomorphic encryption security level;
- Ciphertext dimension is n;
- Ciphertext modulus is q.

Ciphertext dimension n shall be chosen on basis of desired security level and value of ciphertext modulus q. If ciphertext modulus q is bigger than noise budget it enables implementation more complex homomorphic evaluation function f i.e. implementation the function with bigger depth.

Palisade library implements modular arithmetic schemes: BGV and BVF with 128-bit security level beginning from ciphertext dimension $n = 2048$.

The public key encryption operation in BFV scheme has the best performance when the SEAL library is used. Performance difference depends on the ciphertext dimension: while the SEAL encryption is three times faster for the ciphertext dimension of 2048, when the ciphertext dimension is 32768, this factor is 1.3 times. The encryption operation has the best performance in BGV scheme when the Palisade library is used. Performance difference ratio decreases with the increase of the ciphertext dimension. The encryption operation in CKKS scheme for ciphertext dimension $n \geq 8192$ has the best performance when the HELib library is used, whereas in case of lower dimension n the best results are achieved by using SEAL library.

TABLE II
PUBLIC KEY ENCRYPTION

HE scheme	HE parameters		HE library		
	n	$\log_2 q$	Palisade	HELib	SEAL
BFV	1,024	27	-	-	272
BGV	1,024	27	-	1,783	-
CKKS	1,024	27	585	482	257
BFV	2,048	54	1,557	-	506
BGV	2,048	54	1,560	3,608	-
CKKS	2,048	54	1,173	997	479
BFV	4,096	109	3,519	-	1,687
BGV	4,096	109	3,493	7,833	-
CKKS	4,096	109	2,753	2,288	1,926
BFV	8,192	218	7,773	-	4,838
BGV	8,192	218	8,116	17,817	-
CKKS	8,192	218	7,538	4,664	5,688
BFV	16,384	438	24,050	-	16,252
BGV	16,384	438	25,926	44,796	-
CKKS	16,384	438	23,183	12,581	19,344
BFV	32,768	881	77,553	-	59,457
BGV	32,768	881	78,639	109,340	-
CKKS	32,768	881	76,406	39,890	71,373

TABLE III
SECRET KEY DECRYPTION

HE scheme	HE parameters		HE library		
	n	$\log_2 q$	Palisade	HELib	SEAL
BFV	1,024	27	-	-	63
BGV	1,024	27	-	13,047	-
CKKS	1,024	27	415	3,159	10
BFV	2,048	54	159	-	127
BGV	2,048	54	133	49,096	-
CKKS	2,048	54	809	5,104	19
BFV	4,096	109	420	-	416
BGV	4,096	109	353	192,351	-
CKKS	4,096	109	1,432	14,279	72
BFV	8,192	218	940	-	1,484
BGV	8,192	218	1,012	763,178	-
CKKS	8,192	218	6,038	48,960	290
BFV	16,384	438	2,370	-	5,904
BGV	16,384	438	3,690	3,033,690	-
CKKS	16,384	438	13,776	183,254	1,166
BFV	32,768	881	7,330	-	24,919
BGV	32,768	881	14,941	12,003,497	-
CKKS	32,768	881	51,960	701,913	4,826

The decryption operation in CKKS scheme has the best performance by using SEAL library. The decryption operation in CKKS scheme when using SEAL is approximately 10 times faster than when Palisade is used and more than 100 times faster than when HELib is used.

The secret key decryption operation in BGV scheme performs better by several orders of magnitude in the Palisade than in the HELib library.

The decryption operation in BFV scheme for ciphertext dimension $n \geq 8192$ has better performance when Palisade library is used, whereas in case of lower dimension n better results are achieved by using SEAL library.

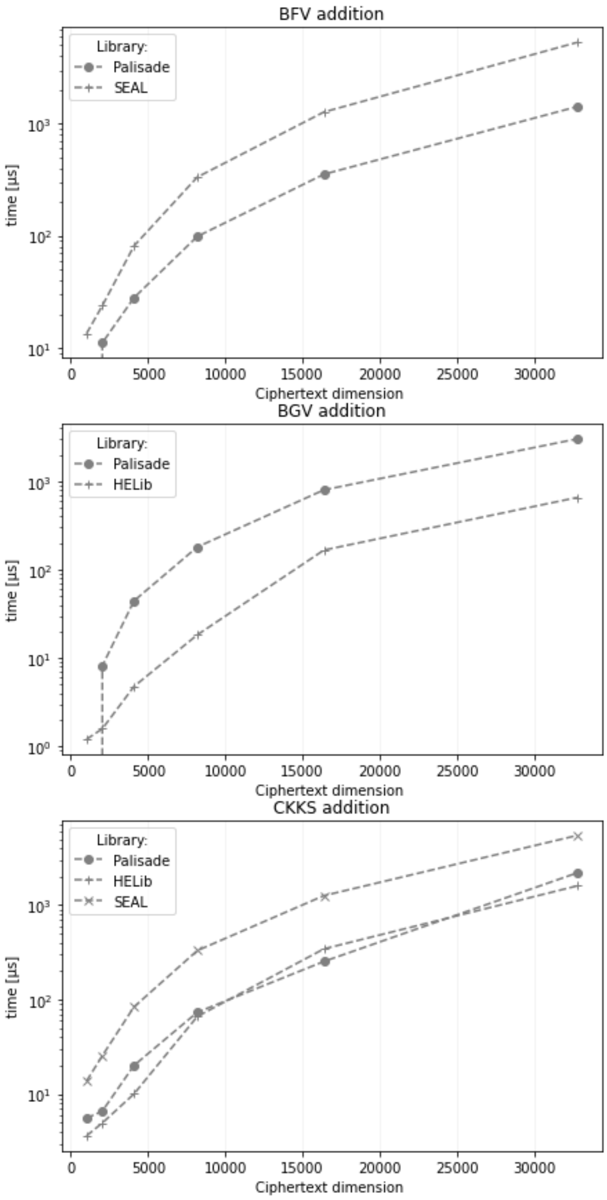


Fig. 4. Homomorphic encryption – addition operation time

The ciphertext addition in CKKS scheme has the best performance in the HELib library. The ciphertext addition in CKKS scheme has better performance in the Palisade than in the SEAL library, but the differences are generally smaller than for the decryption operation.

The ciphertext addition in BFV scheme has significantly better performance (more than 2 times faster) in the Palisade than in the SEAL library.

The ciphertext addition in BGV scheme has significantly better performance (more than 4 times faster) in the Palisade than in the SEAL library.

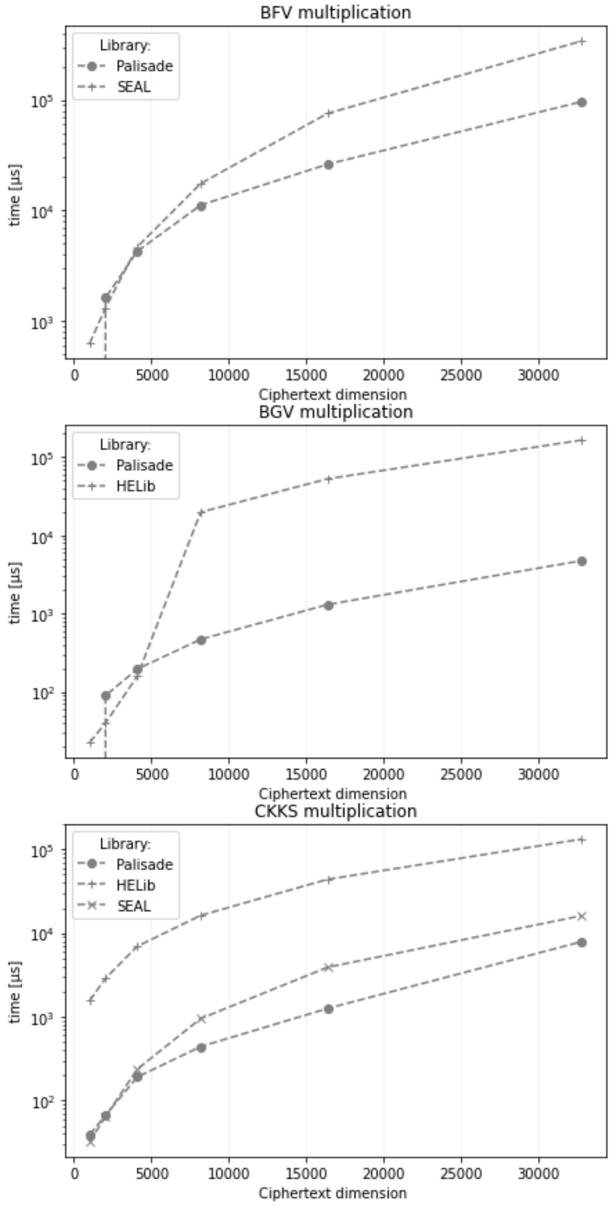


Fig. 5. Homomorphic encryption – multiplication operation time

The ciphertext multiplication is much more complex and more time consuming than ciphertext addition. Figure 4 presents the time needed for performing homomorphic multiplication without relinearization procedure.

The cyphertext multiplication in CKKS scheme for ciphertext dimension $n \geq 8192$ has the best performance when implemented in the Palisade library whereas in case of lower dimension n the better results are achieved using SEAL library.

The cyphertext multiplication in BGV scheme for ciphertext dimension $n \geq 8192$ has significantly better performance (more than 3 times faster) when implemented in the Palisade library than in the HELib whereas for lower ciphertext dimensions better results are achieved by using HELib library.

The cyphertext multiplication in BFV scheme for ciphertext dimension $n \geq 4096$ has better performance in the Palisade than in the SEAL whereas for lower ciphertext dimensions slightly better results are achieved by using SEAL library.

In addition, we compared execution time of homomorphic operations with no security level versus operations with 128-bit security level. We have measured execution time of homomorphic operations in CKKS (approximate arithmetic) and BGV (integer modulo arithmetic) schemes that are implemented in the Palisade library.

In the experiments we have got similar ratio of results for both schemes, so we present only results related to CKKS scheme.

In the tests we have performed homomorphic operations by using following scenarios:

1. No security level with ciphertext dimension n=512;
2. 128-bit security level with ciphertext dimension n=32768.

Each operation was executed 1000 times. In both scenarios it is used same value of ciphertext modulus q.

We have got following results of homomorphic operations (CKKS scheme):

- Public key encryption operation is about 69 times faster in scenario 1;
- Private key decryption operation is about 46 times faster in scenario 1;
- Homomorphic addition operation is about 45 times faster in scenario 1;
- Homomorphic multiplication operation is about 48 times faster in scenario 1.

VI. Conclusions

Homomorphic encryption allows performing computations on the encrypted data, without decrypting them. The paper compares the time needed to execute homomorphic operations, like, public key encryption, secret key decryption, addition and multiplication implemented in the open-source libraries: Microsoft SEAL, Palisade, and HELib. The operations are compared for BGV, BFV and CKKS homomorphic encryption schemes implemented in the libraries.

Homomorphic operations that are performed at client side: public key encryption and secret key decryption if it is used BGV scheme (integer arithmetic) have the best performance when using methods that are implemented Palisade.

Homomorphic operations that are performed at the server side: addition and multiplication are fastest when Palisade library is used for all three tested schemes, except for BGV addition and higher ciphertext dimensions in which cases HELib has slightly better performance.

Execution time of homomorphic operations with no security level versus operations with 128-bit security level was performed and showed that all the operations are still by two orders of magnitude slower than when no security is used which presents an issue when complex machine learning or AI calculations are required.

The performance of current fully homomorphic encryption schemes, especially for large parameters, can still be improved. Further improvement can be achieved by implementation low-level homomorphic operations in an assembly language which is executed on a hardware platform. Also it can be achieved better performance if homomorphic operations are implemented in hardware platforms like Graphics Processing Unit (GPU), Application-Specific Integrated Circuit (ASIC), and Field-Programmable Gate Array (FPGA).

Literature

[1] A. Acar, H. Aksu, A. Selcuk, and M. Conti, "A Survey on Homomorphic Encryption Schemes: Theory and Implementation," ACM Comput. Surv. 1, 1, Article 1, http://dx.doi.org/10.1145/3214303, 2018.

[2] V. Lyubashevsky, C. Peikert, and O. Regev, "On ideal lattices and learning with errors over rings," Journal of the ACM (JACM) 60, no. 6, 2013.

[3] Z. Brakerski, C. Gentry, and V. Vaikuntanathan, "Fully Homomorphic Encryption without Bootstrapping," Cryptology ePrint Archive, Report 2011/277. https://eprint.iacr.org/2011/277, 2011.

[4] J. Fan and F. Vercauteren, "Somewhat practical fully homomorphic encryption," IACR Cryptology ePrint Archive, 2012:144, 2012.

[5] J.H. Cheon, A. Kim, M. Kim, and Y. Song, "Homomorphic encryption for arithmetic of approximate numbers," Cryptology ePrint Archive, Report 2016/421, https://eprint.iacr.org/2016/421, 2016.

[6] A. Viand, P. Jattke, A. Hithnawi, "SoK: Fully Homomorphic Encryption Compilers", IEEE Symposium on Security and Privacy 2021.

[7] C. Aguilar Melchor, M. Kilijian, C. Lefebvre, T. Ricosset, "A Comparison of the Homomorphic Encryption Libraries HElib, SEAL and FV-NFLlib," in: Lanet JL., Toma C. (eds) Innovative Security Solutions for Information Technology and Communications. SECITC 2018. Lecture Notes in Computer Science, vol 11359. Springer, Cham. https://doi.org/10.1007/978-3-030-12942-2_32, 2019.

[8] T. Lepoint, M. Naehrig, "A Comparison of the Homomorphic Encryption Schemes FV and YASHE," in: Pointcheval D., Vergnaud D. (eds) Progress in Cryptology – AFRICACRYPT 2014. AFRICACRYPT 2014. Lecture Notes in Computer Science, vol 8469. Springer, Cham. https://doi.org/10.1007/978-3-319-06734-6_20, 2014.

[9] T. Maha, S. Hajji, and A. Ghazi, "Homomorphic encryption applied to the cloud computing security," in Proceedings of the World Congress Engineering, vol. 1, pp. 4-6, 2012.

[10] L. Ducas and D. Micciancio, "FHEW: bootstrapping homomorphic encryption in less than a second," in E. Oswald and M. Fischlin, editors, Advances in Cryptology – EUROCRYPT 2015 - 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, So_a, Bulgaria, April 26-30, 2015, Proceedings, Part I, volume 9056 of Lecture Notes in Computer Science, pages 617-640. Springer, 2015.

[11] I. Chillotti, N. Gama, M. Georgieva, and M. Izabachene, "Faster packed homomorphic operations and e_cient circuit bootstrapping for tfhe," in Advances in Cryptology-ASIACRYPT 2017: 23rd International Conference on the Theory and Application of Cryptology and Information Security, pages 377-408. Springer, 2017.

[12] O. Regev, "The learning with errors problem," in Blavatnik School of Computer Science, Tel Aviv University Invited survey in CCC, 2010.

[13] J. H. Cheon, K. Han, A. Kim, M. Kim, and Y. Song, "A full rns variant of approximate homomorphic encryption," Cryptology ePrint Archive,Report 2018/931, https://eprint.iacr.org/2018/931, 2018.

[14] Y. Polyakov, K. Rohloff, G.W. Ryan, and D. Cousins, "PALISADE Lattice Cryptography Library User Manual (v1.10.6)", 2020.

[15] S. Sathya, P. Vepakomma, R. Raskar, R. Ramachandra, and S. Bhattacharya, "A Review of Homomorphic Encryption Libraries for Secure Computation," http://arxiv.org/abs/1812.024, 2018.

[16] M. Albrecht, M. Chase, H. Chen and others, "Homomorphic encryption standardization," homomorphicencryption.org, 2018.

[17] K. Laine, "Simple Encrypted Arithmetic Library 2.3.1," 2017.

[18] S. Halevi and V. Shoup, "Algorithms in Helib," in Advances in Cryptology – CRYPTO 2014, J. A. Garay and R. Gennaro, Eds, Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 554–571, 2014.

[19] S.Halevi, V. Shoup, "HElib design principles," 2020.

[20] V. Shoup and others, "NTL: A library for doing number theory," http://www.shoup.net/ntl.

Comparison of Message Queue Technologies for Highly Available Microservices in IoT

Marko Milosavljević, Milica Matić, Neven Jović, Marija Antić

Abstract— **Internet of Things (IoT) solutions connect large numbers of devices, which generate various data and control messages asynchronously. In the IoT system cloud, these messages need to be queued in order to control the processing load and prevent the overload in cases of traffic bursts. On the other hand, one of the requirements the IoT cloud needs to fulfill is the high availability. Therefore, multiple instances of services accepting and processing the messages generated by the devices are needed. There are various message queue technologies available today, but they all have their limitations. In this paper, we compare the performance of Apache Kafka and RabbitMQ in the scenario of the highly available IoT cloud data processing.**

Index Terms— **message queue; high availability; load balancing; internet of things.**

I. Introduction

In the past decade, the world is witnessing the expansion of Internet of Things (IoT) solutions. Within IoT systems, different devices are connected to perform a certain function together. IoT use-cases are various, such as smart transport, smart fabrics, smart cities, smart homes, etc.

In order to collaborate, the devices need to be able to exchange data such as commands and state change reports. Although the expansion of IoT has led to the development of technologies such as ZigBee, Z-Wave, WiFi or Bluetooth Low Energy, which enabled the connection of many different actuators and sensors into large local mesh networks, in order for an IoT solution to achieve its purpose, the existence of the cloud component is also needed. The cloud allows remote control and monitoring of the local networks, but it can also provide advanced features which require processing of larger quantities of historical system data, or the interaction with components responsible for customer management, software update and third-party services.

As the data from the IoT system is generated asynchronously [1], and processing it requires a certain amount of time, mechanisms are needed to control the cloud load. Usually, this control is achieved by deploying various message queueing systems, that allow to communicate between different components of the cloud, and react to messages generated by the end devices [2]. Message queuing technologies which are available today differ in terms of the performance guarantees they offer, and depending on the actual use-case, metrics such as latency, disk space, RAM memory or processor usage may be a limiting factor [2], [3]. The comparison of Kafka and Apache Pulsar has been performed by the authors in [4], and it has been shown that, although Apache Pulsar may achieve better results in terms of resource usage, the maturity of the solution, available documentation, and possibility to integrate with other data processing tools, may be a reason to favor Kafka in the commercial deployment scenarios. On the other hand, Kafka and RabbitMQ have been compared in [5], to show that RabbitMQ has its advantages in terms of the achieved throughput on a single server instance, but the scaling options are on Kafka's side.

In this paper, we explore the possibility of replacing the already implemented RabbitMQ message queueing within the smart home system cloud [6],[7], with Apache Kafka. Within the deployed smart home cloud, messages generated by end devices are processed by multiple cloud services. As the number of supported features is growing, so is the number of the cloud services that process these messages. Also, some of the messages need to be processed by multiple of these services. Additionally, as the number of users grows, the system needs to be scaled up, and, as already said, Kafka has its advantages in this domain. The paper is organized as follows: in Section II, the elements of smart home system and its cloud architecture are introduced, then the overview of RabbitMQ and Kafka is given in Section III and Section IV. Finally, the performed tests and their results are presented in Section V.

II. Smart Home Cloud Data Buffering

In the existing smart home solution, the end devices within the household use technologies such as ZigBee, Z-Wave and ONVIF/IP to connect to the home gateway – Fig. 1. The gateway is responsible to execute the core system logic: it implements the middleware which represents all of the devices in the same way, regardless of the communication technology they use in the local network, and allows them to work together, according to the automation rules set up by the user. To communicate with the user applications and cloud backend, the gateway uses MQTT protocol. MQTT conveys commands issued by the user, system control messages, and reports about device state changes. Control messages are processed on the cloud side, for the purpose of system

Marko Milosavljević is with OBLO Living, Novi Sad, Narodnog fronta 21a, Serbia (e-mail: marko.a.milosavljevic@ obloliving.com).

Milica Matić is with the Faculty of Technical Sciences, University of Novi Sad, Serbia (e-mail: milica.matic@rt-rk.uns.ac.rs).

Neven Jović is with OBLO Living, Novi Sad, Narodnog fronta 21a, Serbia (e-mail: neven.jovic@ obloliving.com).

Marija Antić is with the Faculty of Technical Sciences, University of Novi Sad, Serbia (e-mail: marija.antic@rt-rk.uns.ac.rs).

administration, upgrade, backup and restore. Also, reports about device state changes are stored to provide user with the information about the history of system usage [7].

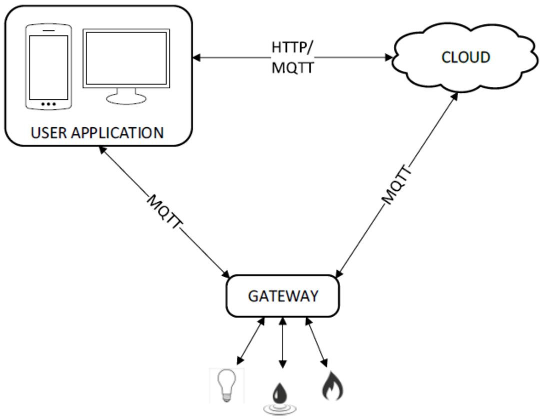


Fig. 1. Smart home system components and communication between them.

The observed smart home cloud system solution has the microservice-based architecture. It is highly available (HA), which means that the entire system is fault tolerant, i.e. that there are multiple instances of every microservice running [6]. In order to prevent problems with MQTT messages processing due to the overload of cloud system, or the failure of some instances, temporary data buffering is necessary. In the temporary data buffering module all important messages are first queued, allowing relevant microservices to process them at their own pace.

In the current implementation, RabbitMQ is used for the purpose of data buffering. The incoming MQTT messages are parsed by the B2Q (Broker to Queue) microservice, and directed to the appropriate RabbitMQ queues, based on the information they contain. All of the instances of one cloud microservice share the load of processing the messages from the RabbitMQ queue they are associated with. The problem here represents the fact that if one message needs to be processed in multiple ways (i.e. it is relevant as the input for multiple cloud microservices), it has to be replicated to multiple queues. Therefore, in this paper we explore the possibility of replacing RabbitMQ with Apache Kafka. We implement the B2K (Broker to Kafka) microservice, which publishes messages to Kafka queues, that the processing microservices are subscribed to, and we compare the performance of the two implementations.

A. RabbitMQ

RabbitMQ is a message queue manager, which has originally implemented the Advanced Message Queuing Protocol (AMQP). Later it was extended to support Streaming Text Oriented Messaging Protocol (STOMP), Message Queue Telemetry Transport (MQTT), and other protocols, but AMQP remains the default and the most widely used one.

RabbitMQ messages can convey any kind of information, from a simple text message to a message with information about processes important for the system. Message broker stores the message into the queue, until the application fetches it for processing. Message queuing allows web servers to avoid the overload, as they can control the number of the messages that are processed simultaneously. It is also useful for distributing messages to multiple consumers sharing the load and providing fault tolerance.

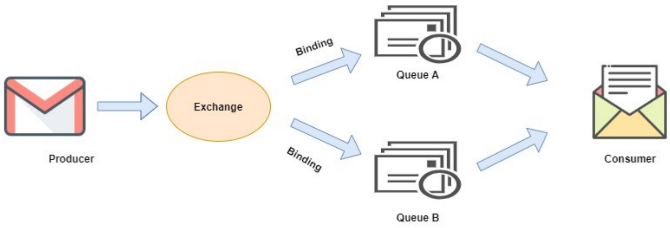


Fig. 2. RabbitMQ message delivery mechanism.

Producer applications create the messages, but the messages are not published directly to a queue. First, the producer sends the message to the RabbitMQ exchange running on the broker – Fig. 2. The exchange is responsible for routing the messages to different queues, based on the configured bindings and routing keys. Four types of exchanges exist - direct, topic, fanout and headers exchange. In the direct exchange, the message is routed to the queue whose binding key matches the routing key of the message. The topic exchange does a wildcard match between the routing key and the routing pattern specified in the binding. The fanout exchange routes messages to all of the queues bound to it. The headers exchange uses the message header attributes for routing. Consumers subscribe to the queues and process the messages from them. All consumers subscribed to the same queue will share the load of processing the messages from that queue. The messages are deleted from the queue after processing.

B. Apache Kafka

Apache Kafka is an event streaming platform. It is elastic, distributed, highly scalable and fault-tolerant. Similar to RabbitMQ, Kafka has the client and server side. Kafka clients and servers communicate using TCP protocol.

Kafka implements the publish/subscribe mechanism, and allows processing streams of events as they arrive into the system or retrospectively, but also allow to store streams of events as long as they are needed.

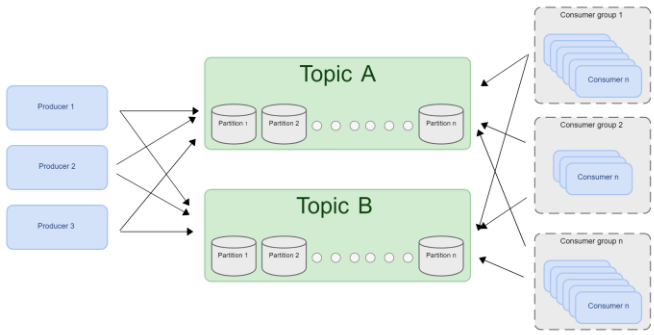


Fig. 3. Kafka message processing mechanism.

Similar to RabbitMQ, the Apache Kafka clients can act as producers and consumers – Fig. 3. Producers represent client applications that write (publish) events to Kafka. On the other hand, consumers are subscribing to topics, reading and writing events. Producers and consumers are not aware of each other. They work completely independently, and that is a key design to achieve high scalability. Therefore, producers will never need to wait for consumers.

When data is written to Kafka, it is written in the form of an event containing the key, value, timestamp and optional metadata. Events are stored in topics. The durability of events inside Kafka's topic is configurable. Unlike RabbitMQ, Kafka events can be read whenever they are needed, because events are not deleted after consumptions. Events can be stored as long as needed. Storing data for a long time does not affect Kafka.

Topics in Kafka are partitioned, and one Kafka topic can have any number of partitions defined in the Kafka configuration file. Events are ordered inside the partition in the exactly same order as they were written, and one consumer can process data from one partition only. However, the data stored in one partition can be processed by multiple consumers belonging to different consumer groups, i.e. one message can be processed multiple times, without the need to duplicate it. Offset is an integer number that is used to maintain the current position of a consumer inside partition. Every topic can be replicated, so that there are dozens of brokers that have a copy of data. This makes data fault-tolerant and highly-available.

III. Testing and Results

Tests were designed to measure CPU load of smart home system servers when RabbitMQ and Apache Kafka are used for data buffering. RabbitMQ and Kafka brokers were run on the 8-core Intel i7 processor with 8 GB of RAM memory.

TABLE I
RabbitMQ test results

Setup	CPU usage on 8 cores [%]		
	average	**maximum**	**deviation**
16 producers 8 queues 0 consumers	324	640	108
16 producers 8 queues 8 consumers	410	794	197
16 producers 8 queues 16 consumers	486	800	167
16 producers 8 queues 32 consumers	553	800	147

To test the RabbitMQ buffering, 16 producer B2Q processes were created, that published messages to 8 queues. The messages from these queues were processed by a variable number of consumers (0, 8, 16, 32). Producers were configured to publish messages every 1 ms. Test results are presented in Table I.

RabbitMQ reached CPU limit after 16 consumers, but was able to continue working stably, while the setup with 32 consumers stopped working after ten minutes. The throughput of the system was approximately 11000 messages per second. Maximum CPU usage was 800%, i.e. all eight cores were used 100%.

To test Kafka performance, 16 producers were created, which published to the variable number of partitions (32, 64, 128). Since Kafka allows only one consumer per partition, the number of consumers was also varied from 0 to 128. Test results are presented in Table II.

In any of test cases limit of Kafka maximum CPU load was not reached. It can be observed that the CPU usage deviation is smaller than in RabbitMQ case. Therefore, the server stays stable, even as the number of messages that are stored in Kafka increases with time.

TABLE II
Kafka test results

Setup	CPU usage on 8 cores [%]		
	average	**maximum**	**deviation**
16 producers 32 partitions 0 consumers	210	573	65
16 producers 32 partitions 32 onsumers	202	347	43
16 producers 64 partitions 0 consumers	186	473	90
16 producers 64 partitions 64 consumers	208	360	37
16 producers 128 partitions 0 consumers	150	300	93
16 producers 128 partitions 128 consumers	480	553	53

IV. Conclusion

This paper gave a brief description of some of the message queueing technologies that can be used for flow control and load balancing in the IoT scenario. RabbitMQ and Apache Kafka were deployed within the smart home system cloud, and their performance was tested for a variable number of consumers.

The presented test results indicate that data buffering in Kafka is highly stable and has the lower average CPU usage. At any point of testing, maximum CPU usage was never reached. Therefore, in our further work we will focus on integrating Kafka in the data collection and storage module of the smart home system. Using Kafka will allow us to process the same messages multiple times, without the need to duplicate data. This, in turn, opens the possibility to create advanced data processing scenarios which may bring added value to the users of the smart home system.

V. Acknowledgment

This research (paper) has been supported by the Ministry of Education, Science and Technological Development through the project no. 451-03-68/2020-14/200156: "Innovative scientific and artistic research from the FTS (activity) domain".

References

[1] F. Metzger, T. Hoßfeld, A. Bauer, S. Kounev and P. E. Heegaard, "Modeling of Aggregated IoT Traffic and Its Application to an IoT Cloud," *Proceedings of the IEEE*, vol. 107, no. 4, pp. 679-694, April 2019

[2] G. Fu, Y. Zhang and G. Yu, "A Fair Comparison of Message Queuing Systems," *IEEE Access*, vol. 9, pp. 421-432, Jan. 2021

[3] H. Wu, Z. Shang and K. Wolter, "Performance Prediction for the Apache Kafka Messaging System," *Proc. of IEEE HPCC/SmartCity/DSS*, Aug. 2019

[4] S. Intorruk and T. Numnonda, "A Comparative Study on Performance and Resource Utilization of Real-time Distributed Messaging Systems for Big Data," *Proc. of IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, July 2019

[5] P. Dobbelaere and K. S. Esmaili, "Kafka versus RabbitMQ: A comparative study of two industry reference publish/subscribe implementations: Industry Paper," *Proceedings of the 11th ACM International Conference on Distributed and Event-based Systems (DEBS '17)*, June 2017

[6] M. Matić, E. Nan, M. Antić, S. Ivanović and R. Pavlović, "Model-Based Load Testing in the IoT System," *Proc. of International Conference on Consumer Electronics (ICCE-Berlin)*, Sept. 2019

[7] S. Ivanović, M. Antić, I. Papp, N. Jović, "Data Acquisition, Collection and Storage in Smart Home Solutions," *Proc. of 6th International Conference on Electrical, Electronic and Computing Engineering (IcETRAN)*, May 2019

Design of a Network Topology Using CISCO NSO Orchestrator

Mioljub Jovanovic
Wireless Communications Research Group
University of Westminster
London, UK.
Cisco Systems, Diegem, Belgium
m.jovanovic@my.westminster.ac.uk
mjovanov@gmail.com

Milan Cabarkapa
Department of Telecommunications
School of Electrical Engineering
University of Belgrade
Belgrade, Serbia.
cabmilan@etf.rs

Djuradj Budimir
Wireless Communications Research Group
University of Westminster
London, UK.
School of Electrical Engineeing
Univrsity of Belgrade, Serbia.
d.budimir@wmin.ac.uk
d.budimir@etf.rs

***Abstract*— This paper presents the design of a network topology using CISCO NSO orchestrator. The mismatch problem solution between a network service and its monitoring is proposed. Applying the proposed approach, the telemetry efficiency ratio parameter greater than 40 is achieved. All tests are performed in the real experimental conditions using CISCO NSO orchestrator.**

Keywords—intent based network; intent-aware monitoring agent; model-driven telemetry; service assurance

I. Introduction

NFV orchestrators (e.g., Tacker [1], Cloudify [2], ONAP [3], CISCO NSO [4]) are a crucial part for the dynamic and optimal management and orchestration of various virtualized network resources (e.g., VMs, Virtualized Network Functions). 5G technology, empowered by NFV and SDN, presents a new dimension of complexity that must be addressed by service assurance [5].

Using these orchestration software having higher level of abstraction, the rapid connectivity and provisioning could be achieved at lower prices while letting to operators possibility to build, arrange and preserve network service [6], [7].

Communication Service Provider (CSP) networks – such as Virtual Evolved Packet Core are subject to very dynamic configuration change. Provisioning, modification and termination of packet data services are being done in rapid pace in order to keep up with dynamic environment needs and cater to main business drivers, such as IoT, Video etc. SDN technologies using Network Slicing approach are foundation for such a dynamic environment, allowing automated and programmatic configuration of network services [5].

Traditionally network services are being monitored by deployment of probes which generate traffic and provide feedback on the status of the service. Due to such rapid changes in network service configuration there is open question in regard to monitoring and assuring provisioned services: What is the right approach to take in order to monitor the network which constantly changes? How to ensure network service is operational and carefully selecting probes to monitor network service? [5], [8].

Monitoring using active probes face challenges such as introduction of synthetized traffic within the data flow, end to end monitoring only with no understanding of the data path, lack of comprehension of the configuration intent etc [9]-[11]. Generated traffic using probes should resemble real traffic of the network service, however even with almost perfect synthetized traffic, there is substantial possibility that real network service traffic could be impacted, but probe does not detect such a problem since probe is not part of the actual real data flow [12]. Therefore, there is a gap in regard to monitoring and assurance of the actual network service data flow, with all network elements data traverses on the path between endpoints.

We are proposing solution based on Intent Based Networking (IBN). The proposed approach consists by:

- Extraction of configuration intent by analysing of the network service configuration.
- Discovery of the network elements along the network service data path.
- Leveraging existing network monitoring capabilities of network elements, along with probes and Model Driven Telemetry (MDT) to get more accurate information on the status of desired Network Service.

Research methodology used in understanding benefits of proposed monitoring approach involves qualitative approach, comparative analysis of existing - probe based monitoring and proposed solution based on Intent Based Networking (IBN). By using the proposed approach, we have achieved higher than 40 for the telemetry efficiency ratio parameter.

II. Intent Based Network Methodology

Role of Intent Based Network is transforming Business Intents into configuration changes. As depicted in Fig. 1, Intent at high level represents one or set of different requirements which describe service or network.

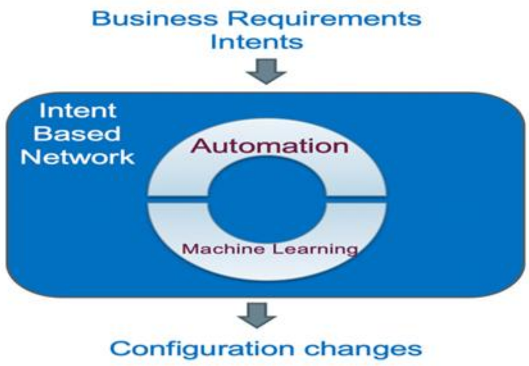


Fig. 1. Intent Based Network – high level description

Those requirements are then being analysed by set of steps, processes or algorithms in order to convert/render high-level requirements into lower form of abstraction, which could then be used to configure computer network elements in order to enable needed service. As business intent is being transformed into configuration on devices it's important to enable monitoring of the network services in order to have understanding whether desired service is operational and functioning in accordance to the business requirements – key performance indicators (KPIs). As graphically demonstrated in Fig. 2, traditionally in legacy network such monitoring would mean enabling monitoring on different data points including but not limited to: SNMP, Netflow/SFlow, telemetry and even Command Line Interface outputs (CLI). Acquiring data from different sources would certainly improve visibility on the state of the network, yet it would greatly impact efficiency and would aplify amount of telemetry data transferred over the network, but without providing clear answer on whether the intent has been fulfilled and whether network service is running and operational as per pre-defined KPIs [13].

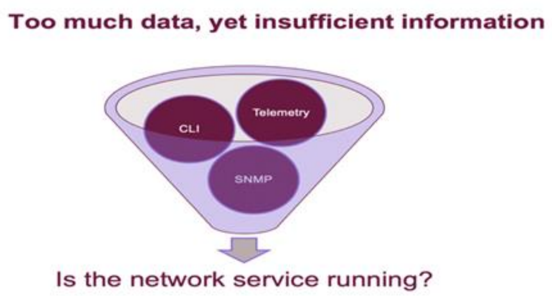


Fig. 2. Main query is - Is the Network Service running according to the pre-defined KPIs?

III. Experimental Setup

Experimental setup consists of the following routers: Simulated customer premises routers (CE), provider core routers (P) and provider edge routers (PE). Fig.3, shows the network with service models which is configured using the orchestration network architecture.

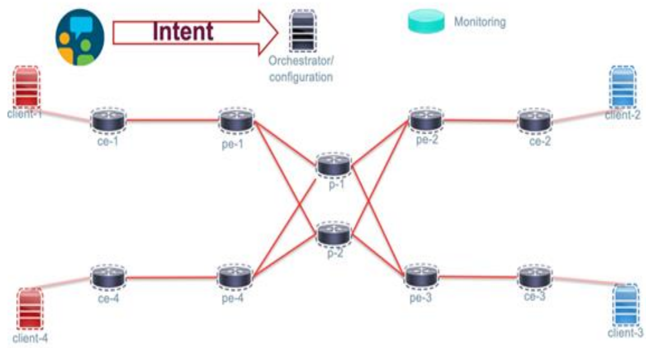


Fig. 3. Business Intent communicated to the orchestrator

In Fig. 4 orchestrator is configuring devices in order to fulfil desired service intent. Orchestrator uses Netconf protocol to access and configure network elements which are taking part in the data path to enable desired service.

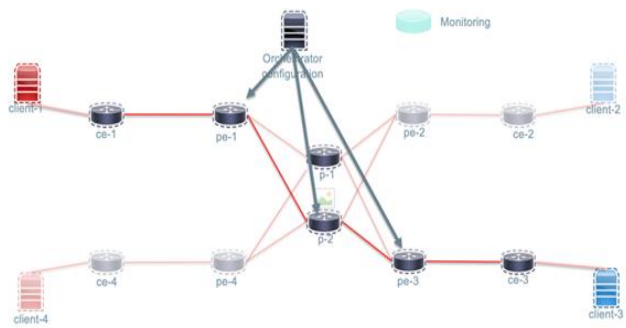


Fig. 4. Orchestrator sends configuration to network devices

In the provided example, actual intent is to establish communication – tunnel service between ce-1 and ce-3 network device in order to enable communication between Client-1 and Client-3. In order to traverse path between Client-1 and Client-3, data packets need to cross pe-1, p-2 and pe-3 as shortest path between the endpoints. Of course, this trajectory may be different in function of routing protocols and connectivity in function of time, but topology discovery and update events will be discussed in future work. At this time, we are focusing on the fixed path through the experimental network and assuming there would not be topology changes throughout the experiment shown in Fig. 5.

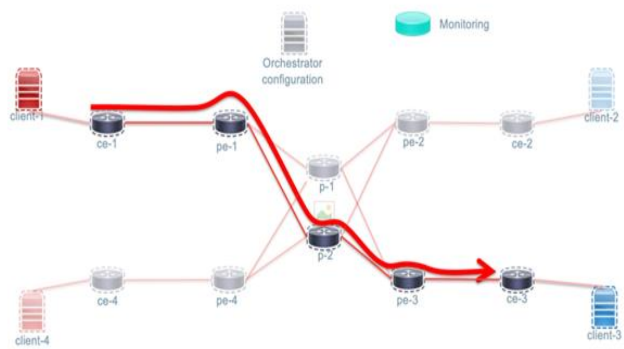


Fig. 5. Service is configured. Question: Service running within acceptable KPIs? Question: Is configuration model mapped to monitoring model?

In Fig. 6 we can observe each of the network devices streaming telemetry data to the collector, monitoring platform which is receiving and processing all telemetry data.

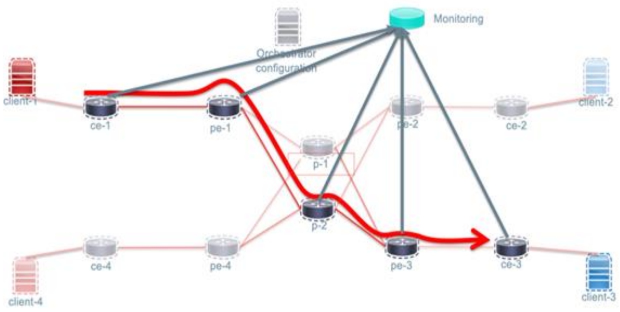


Fig. 6. Telemetry data streamed to Monitoring/Analytics platform. 250000 different stats per router (740 kbps of data)

Thanks to the fact involved network elements are already using Model Driven Telemetry processing data points by collector is simpler. However, as there are so many different data points which are being monitored on devices, there may be information overload since on average router there could easily be 250000 different monitored data points. Such as large number of collected data points could essentially mean that amount of generated telemetry data may be significantly high and could pose challenge for network infrastructure as well as could cause impact to collector processing capacity.

Instead of monitoring all relevant and non-relevant data points, causing unnecessary increase of traffic and compute resources to process large amount of data, we're proposing significant reduction in amount of telemetry data by ensuring that only minimal set of relevant data points is exported from the network devices by means of intent-aware monitoring agent (IAMA). Data reduction task is accomplished by deploying IAMA locally to the network devices, thus leveraging local area network (LAN) links and avoiding use of wide-are links (WAN) for large amount of data points. IAMA is aware of the service details and is also capable of receiving telemetry data. As represented on IAMA architecture in Fig. 7, service intent is received by from the orchestrator while MDT is received from network devices.

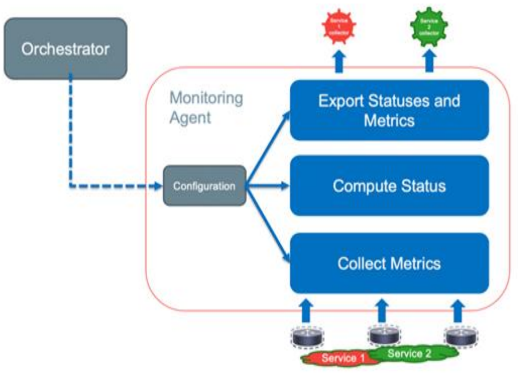


Fig. 7. Intent-aware monitoring agent architecture

IAMA is performing analysis on the received datasets and series of computations in order to determine actual state of the service. Steps performed by IAMA: collecting MDT, processing and exporting reduced – yet more relevant MDT is called IAMA pipeline. Final result of IAMA pipeline is significantly reduced amount of MDT containing only high-level status of the monitored service, as per pre-defined Key-Performance Indicators (KPIs).

IV. Results

Measuring objective was to determine how much data is actually received via MDT under usual telemetry export, with typical data points for router such as environmental, interface stats etc. Result of this work outlines amount of measured data after performing analysis of the incoming telemetry and mapping to service aware MDT. All routers and all incoming data points were taken into account.

TABLE I. EXPERIMENTAL RESULTS

	Intent-Aware Monitoring Efficiency			
	Total MB	**Rate 1 min in kbps**	**Rate 5 min in kbps**	**Rate 15 min in kbps**
Incoming from routers	5200	740.7	700.1	711.7
This work	130.8	17.3	17.2	17.1
Outgoing to Analytics platform	224.9	29.5	29.1	29.3
This work efficiency ratio	40.9	42.8	40.6	41.7

As outlined in Table I, demonstrated experimental results have reduced the amount of incoming MDT from routers from 5.2 GB to 130 MB, while preserving relevant information which is – is service running and operational per pre-defined KPIs.

V. Conclusion

The design of a network topology using CISCO NSO orchestrator has been presented in this paper. The solution about mismatch problem between a network service and its monitoring has been proposed. The telemetry efficiency ratio parameter of more than 40 has been achieved. The amount of telemetry data has been reduced by injecting service aware information in MDT and removing all overhead MDT data points which do not need to be exposed to the network operator who is monitoring the service. Of course, full MDT can also be enabled if desired.

References

[1] https://docs.openstack.org/tacker/latest/

[2] https://cloudify.co/

[3] https://www.onap.org/

[4] https://www.cisco.com/c/en/us/solutions/service-provider/solutions-cloud-providers/network-services-orchestrator-solutions.html

[5] Anil Rao, "Reimagining service assurance for NFV, SDN and 5G", White paper, Analysis Mason, 2018.

[6] R. Mijumbi, J. Serrat, J. l. Gorricho, S. Latre, M. Charalambides, and D. Lopez, "Management and Orchestration Challenges in Network Functions Virtualization," IEEE Communications Magazine,vol. 54, no. 1, pp. 98–105, Jan 2016.

[7] A. J. Gonzalez, G. Nencioni, A. Kamisiski, B. E. Helvik, and P. E. Heegaard, "Dependability of the NFV Orchestrator: State of the Art and Research Challenges," IEEE Communications Surveys Tutorials, pp. 1–23, 2018.

[8] M. Pattaranantakul, R. He, Z. Zhang, A. Meddahi and P. Wang, "Leveraging Network Functions Virtualization Orchestrators to Achieve Software-Defined Access Control in the Clouds," in IEEE

Transactions on Dependable and Secure Computing, pp. 1-14, Nov. 2018.

[9] A. D'Alconzo, I. Drago, A. Morichetta, M. Mellia and P. Casas, "A Survey on Big Data for Network Traffic Monitoring and Analysis," in IEEE Transactions on Network and Service Management, vol. 16, no. 3, pp. 800-813, Sept. 2019.

[10] R. Boutaba, M. A. Salahuddin, N. Limam, S. Ayoubi, N. Shahriar, F. Estrada-Solano, and O. M. Caicedo. A Comprehensive Survey on Machine Learning for Networking: Evolution, Applications and Research Opportunities. J. Internet Serv. Appl., 9(16), 2018.

[11] Cisco Systems, Inc, "GitHub Network Telemetry Pipeline," Cisco Systems, Inc, 2017. [Online]. Available: https://github.com/cisco/bigmuddy-network-telemetry-pipeline

[12] M. Jovanović, M. Čabarkapa, B. Clause, N. Nešković, M. Prokin, B. Đurađ, Model driven telemetry using Yang for next generation network applications, 5th International Conference on Electrical, Electronic and Computing Engineering (IcETRAN) 2018, pp. 1186 - 1189, Palić, Serbia, June, 2018.

[13] B. Claise, J. Clarke, and J. Lindblad "Network Programmability with YANG: The Structure of Network Automation with YANG, NETCONF, RESTCONF, and gNMI", Addison-Wesley Book, 1st edition, 2019.

Visualization of microscopic morphological characteristics used for determination of infectious molds

Mina Milanović, Aleksandar Milosavljević and Marina Ranđelović

Abstract—**Invasive fungal infections (IFI) and systemic fungal infections (SFI), caused by molds are on the rise, based on data from literature. Diagnostics of those infections can sometimes be inefficient; they require a longer period of time in laboratory procedures and sometimes may lead to late diagnosis or misdiagnosis, which can result in patient's critical condition or even mortality. The goal of this research is to develop a neural network model that will perform identification of molds, and thus accelerate the process of diagnostics. A classifier has been developed, using an EfficientNet-B1 deep convolutional neural network (CNN) and sample images obtained at the Department of Microbiology and Immunology, Medical faculty, University of Niš, Serbia, archives. We applied Grad-CAM visualization to determine morphological characteristics used by the model to classify samples.**

Index Terms—**molds identification, fungal infection, convolutional neural networks, deep learning, Grad-CAM.**

I. Introduction

Ability of fungus to start a pathological process in the host organism is as a specific phenomenon, according to numerous authors, because, excluding groups of dermatophyte molds and tropical fungi, these microorganisms does not need pathogenicity for their dissemination and survival in nature [1]. Among 400.000 species of fungi known in the nature, around 50 kinds can cause invasive fungal infections(IFI), that are characterized by very high morbidity (serious clinical case) and mortality. Numerous reasons have contributed to the increase of number of infections among humans, and incidences of IFI caused by molds are constantly growing. The most important reasons are complex procedures and medical interventions, intensive treatments with antibacterial drugs, cytostatics, immunosuppressants; longer lifespan of a humans, increase in the number of patients at high risk due to primary diseases and treatment, the appearance of resistance in fungi and certainly the establishment of mycological analyzes and higher diagnostic efficiency, i.e. more successful diagnostic procedures in a microbiology [2].

Mina Milanović – Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia (e-mail: mina.milanovic@elfak.rs).

Aleksandar Milosavljević – Faculty of Electronic Engineering, University od Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia (e-mail: aleksandar.milosavljevic@elfak.ni.ac.rs).

Marina Ranđelović - Department of Microbiology and Immunology, Medical faculty, University of Niš, Blvd Zorana Djindjica 81, 18000 Niš, Serbia (e-mail: marina87nis@gmail.com).

Fungi are eukaryotic microorganisms. In nature, they are widespread, living in soil and water, on organic materials as saprophytes, as symbionts or parasites of animals, plants, or human [3]. Based on the structure, all fungi can be primarily divided into yeasts -unicellular fungi with basal cell blastoconidia (blastospora) and multicellular fungi (molds) with a basic hypha cell. Molds classification is performed on the basis of structure, i.e. macroscopic and microscopic morphological characteristics. Differences between the morphology of molds, hypha structure, production of different conidiae (spores), enable a diagnostic procedure for their identification [4].

This goal of this project is to develop a neural network model that will perform identification of molds, and thus accelerate the process of diagnostics. No similar projects, involving determination of molds or their morphological characteristics, could been found during our research, and beside some rapid tests that can be used only for most common types of infections, whole process of determination is manual and sometimes takes days, so providing an application that can accelerate the process can be very beneficial. During recent years, number of infections caused by more rare species of fungi has drastically increased, which was a motivation for a project like this, which includes classification of so far neglected types of fungi.

Sample collection has contained high resolution images, which needed manual preparation for training, as described in chapter II. Prepared dataset was expanded before training and EfficiencyNet-B1 architecture of convolutional neural network (CNN) has been used for developing and training the model, which makes the core of the classifier, as presented in chapter III. Results and discussion of the results, including visualization of decision making process using Grad-CAM method, have been shown in chapter IV. Conclusion and planned further steps have been described in chapter V.

II. Data

A. Dataset description

Fungi, based on morphology are classified in group of yeasts -unicellular fungi with basal cell blastoconidia (blastospora) and multicellular fungi (molds) with a basic hypha cell. Molds can be primarily divided into dermatophytic and non-dermatophytic fungi [5].

Dermatophytic molds, in other words, dermatophytes, are causative agents of superficial fungal infection of skin, hair and nail with prevalence of 22-25% worldwide.

Other group of molds caused invasive fungal infection (IFI) and in recent years incidence of these diseases has been on the rise [6].

Diagnostics of infections caused by dermatophytic and non-dermatophytic fungi can sometimes be inefficient; they require a longer period of time in laboratory procedures and sometimes may lead to late diagnosis. In case of SFI (systemic fungal infection) late diagnosis or misdiagnosis can lead to wrong treatment, as well as not implementing measures for preventing the spread of infection [7]. On the other hand late diagnosis or misdiagnosis of IFI can result in patient's condition impairment or even mortality. In our previous paper [8], we considered only fungi genera that cause invasive infections, but we extended the dataset so types of fungi that cause systematic fungal infections are included too.

In both groups molds classification was performed on the basis of structure, i.e. macroscopic and microscopic morphological characteristics. Expert's knowledge and experience are needed for differentiation and identification of isolated fungi in laboratory practice.

B. *Morphologial differences of fungi*

Microscopic morphological characteristics of Dermatophytes are:

i) ***Microsporum*** spp. are characterized by segmented hyphae, numerous macroconidiae that are thick walled, rough, present microconidia;

ii) ***Trichophyton*** spp. are characterized by segmented hyphae, rare macroconidiae that are thin walled and smooth, numerous microconidiae;

iii) ***Epidermophyton*** spp. are characterized by segmented hyphae, numerous macroconidiae that are thin and thick walled, smooth and microconidiae are not formed.

Microscopic morphology of non-dermatophytic genera is characterized by:

i) ***Aspergillus*** spp.: Septate hyphae with unbranced conidiophores which ending with swollen vesicule that is covered with flak-shaped phialides on which are chains of mostly round sometimes rough conidia;

ii) ***Penicillium*** spp.: Septate hyphae with branched or unbrenched conidiophores that have secondary branches known as metulae or prophialides on which are phialides with chains of conidiae (Figure 1);

iii) ***Fusarium*** spp.: Septate hyphae with formation of canoe shaped or sickle shaped multiseptate macroconidia that are produced from phialides on unbranched or branched conidiophores;

iv) ***Alternaria*** spp.: Septate, dark hyphae with septate conidiophpores and formation of large macroconidiae which have transverse and longitudinal septations;

v) ***Mucor*** spp.; Wide and practically non-septate hyphae, speorangiophores are long, often branched and bear terminal round spore-filled sporangia.

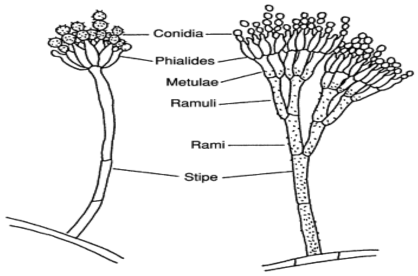


Figure 1. Penicillium morphology

C. *Preparation of dataset images for training*

Machine processes a picture as an array of pixels and numbers, so classification of images can be a rather difficult job, especially in cases where brightness is not the best, position of camera changes or the object is not fully present on the picture, which doesn't present a problem for a person. But, like a human, machine learns in the same manner, with examples of different categories with labels, so it eventually can recognize the patterns on the images.

For our model, we extracted examples of eight fungal genera, which are *Aspergillus* spp., *Fusarium* spp., *Epidermophyton* spp., *Alternaria* spp., *Microsporum*spp., *Penicillium*spp., *Trichophyton*spp. and *Mucorales* spp. (Figure 2). Images have been made at the Department of Microbiology and Immunology, Medical faculty, University of Niš, Serbia, laboratories, where molds have been isolated from patient materials, examined on microscopes and then photographed.

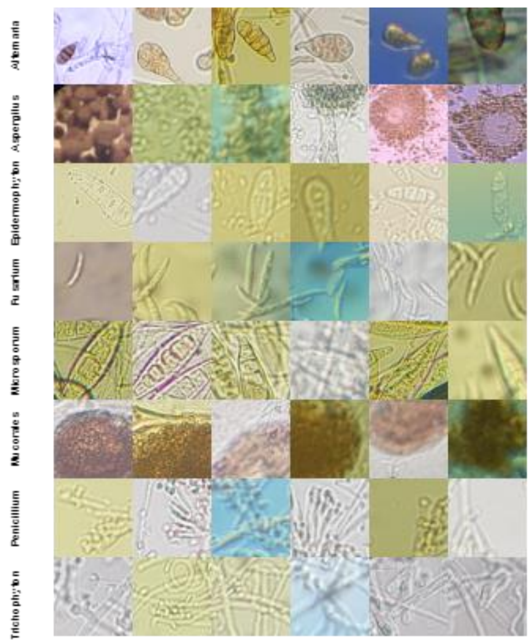


Figure 2. Examples of the dataset images

After preparing the images, which includes manually cutting the high resolution (3024 x 4032 pixels) sample images obtained from Department of Microbiology and Immunology and selecting ones which contain significant molds parts, it is necessary to determine which percentage of them will be used for training, and which for evaluation, since these sets have to be different so results of evaluation can be regular. After manual preparation, there were 6918 images, from which we used around 80% for training and the rest of the images (20%) for evaluation. In Table I, details of dataset used for training are presented.

TABLE I
Details of used dataset

Number of classes	Number of samples	Number od samples per class	Number of images after preparation	Images used for training	Images used for validation
8	492	50-65	6918	5603	1315

III. METHOD DESCRIPTION

For a neural network to learn to recognize certain patterns in images, it is necessary to create examples so it can learn from them. Sample images of patient materials with molds are high resolution, taken on microscopes, and they have to be cut, because of the GPU limitations when it comes to neural network training, and also to make more examples for network to learn. To obtain small resolution images, it was necessary to cut original images into the set of smaller images, suitable for training. After cutting the images, and manually eliminating the ones that don't contain mold patterns, it was decided to expand the dataset so examples can be more informative.

Operations that are used on the images to widen the dataset and provide multiple examples from one image are called augmentations [9]. Using different brightness, rotation, translation, flipping of the images, etc., we made more examples for training (Figure 3). In the end of this process, dataset became more informative and training could be started.

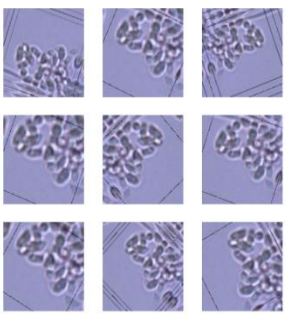


Figure 3. Augmentation of an image gives more images for training

Image classification is a very common problem, present in many different fields of expertise, and traditional approach to this problem is crafting a feature extractor that can be used for training a classifier [10-13]. Earlier solutions used artificial neural networks (ANNs) [14], but major advantages in this area have been made in recent years with development of convolutional neural networks (CNNs) [15]. CNNs represent an aggregation of three architectural ideas, local receptive fields, shared weights and spatial subsampling, which makes them more consistent in terms of translation and distortion [16].

During recent years, many different types of convolutional neural network architectures have been developed, but the one that gave the best result while training our model is EfficientNet. EfficientNet has a family of models (B0 to B7) and during training we tried various variants, where B1 showed the best results, based on accuracy measured. This models, introduced in 2019, by Tan and Le [17], are among the most efficient models, and their innovation lays in heuristic way to scale the model (compound scaling), making them a good combination of efficiency and accuracy [18].

Unlike conventional scaling methods (b-d on Figure 4) that arbitrary scale a single dimension of the network, compound scaling method uniformly scales up all dimensions. In this method, appropriate scaling coefficients are determined with grid search, which discovers relationships between different scaling dimensions. Applying those coefficients to baseline network gets the desired target model size [19].

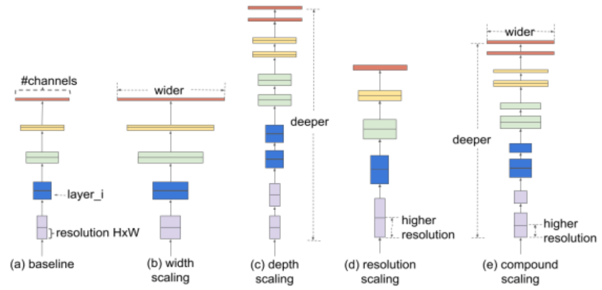


Figure 4. Comparison of different scaling methods [17]

Programming language Python [20] and library Keras have been used for training the model. Keras library [21], implemented in Python, has an interface which can be used for creating and training neural network models, including EfficientNet family. Keras is a deep learning API, running on top of the machine learning platform TensorFlow [22]. They were developed with a focus on enabling fast experimentation.

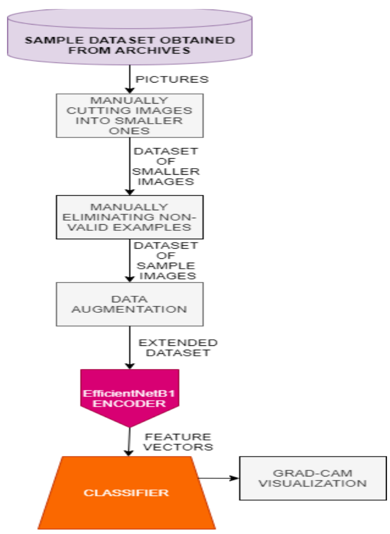


Figure 5. Solution diagram

Model has been compiled with ***RMSprop*** algorithm [23] for optimization (optimizers module),

sparse_categorical_crossentropy type of error (losses module), and the only parameter of metric during learning has been set as accuracy.

EfficientNet-B1 architecture model makes the core of this solution. After training of this model, feature vectors are obtained, which are then used to form a classifier. Classifier can then be used to determine which of 8 classes of molds new input images belong to. Diagram of current solution is shown in Figure 5.

Adjusting parameters of Keras functions and starting the training with different number of epochs, results at these phase of the project show that the trained model after twenty one epochs gives the best results, with 95,74% validation accuracy in classification of images (Figure 6). In our previous paper [8], with a slightly different (including only invasive fungi infections) and drastically smaller dataset, we got the accuracy of around 92%, which shows that we reached a very good improvement with new model. Also, in our previous work we haven't tried EfficientNet neural networks, which gave the best accuracy for our, now expanded, dataset.

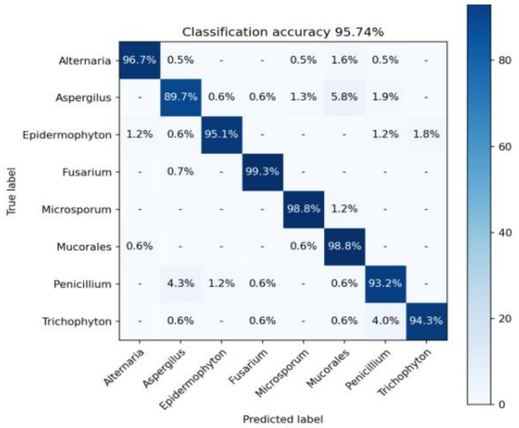


Figure 6. Confusion matrix showing accuracy in %

IV. RESULTS AND DISCUSSION

Because of specific nature of the dataset and sample making, model has not been compared and tested with other datasets or models. In Table II, average results for each fungi genera have been presented.

TABLE II
Results for different fungi genera

Fungi genera spp.	Accuracy [%]	Samples placed correctly/samples per class
Alternaria	96,7	177/183
Aspergillus	89,7	139/155
Epidermophyton	95,1	155/163
Fusarium	99,3	144/145
Microsporum	98,8	166/168
Mucorales	98,8	161/163
Penicillium	93,2	151/162
Trichophyton	94,3	166/176

After validation of the model, it has also been tested manually, showing that the results for most images are accurate. Figure 7 shows confusion matrix, which contains accuracy results per classification class, showing problematic areas too. The most misclassifications happened for *Apergillus* spp. genera, for which we had the least number of clear images, which points out that more images have to be obtained or existing images should be sharpened, so better accuracy can be achieved.

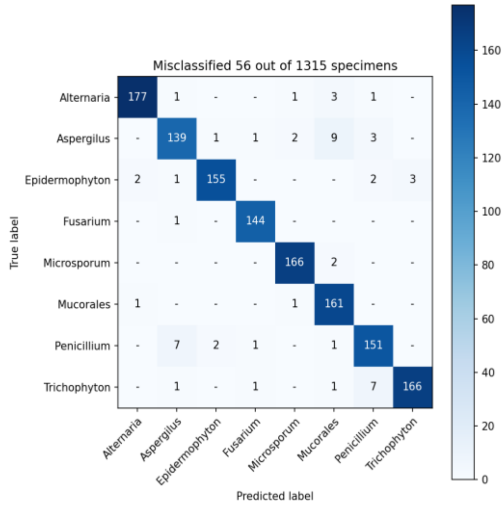


Figure 7. Confusion matrix

Taking into consideration that neural networks learn from examples, from which they learn patterns, and that some sample molds images contain not only significant parts used for diagnostics, but also other parts of materials (for example plain parts of the branches, end of slides on the microscope, different base colors) it is important to verify those learned patterns to be sure that classification, and later diagnostics, performed by the model is valid.

Grad-CAM method is a technique used for visualization of decisions from CNN models, making the decision making process transparent and understandable [24]. This method uses gradients of a target concept (in our cases molds) flowing into final convolutional layer in a network, so it can highlight regions of significance. This way, part of the image which had lead to decision of the classifier is highlighted.

Based on the majority of heat maps got from Grad-CAM method, decisions made by our classier have been done on significant parts of mold samples. Figure 8 shows the examples.

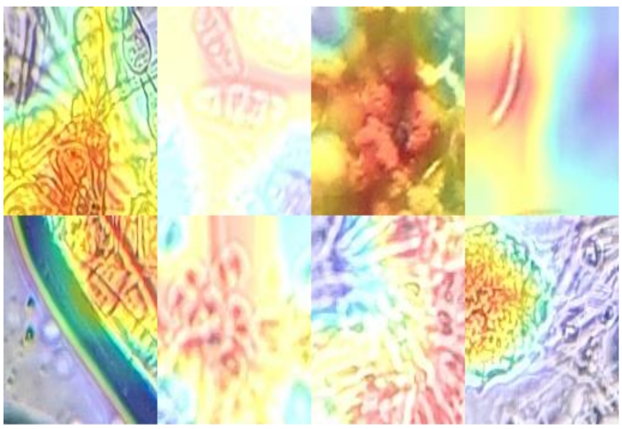


Figure 8. Examples of good pattern recognition visualization

Grad-CAM method is very useful in terms of concluding which of the test images have been misclassified because of the wrong pattern recognition in wrong part of the image (Figure 9). In our case, most of the misclassification happened because of poor quality of input images, because some of them are taken by mobile phones brought close to the microscope oculus, which can result in blurry image. In this way, visualizing the decision making process pointed out that maybe images should be sharpened before processing.

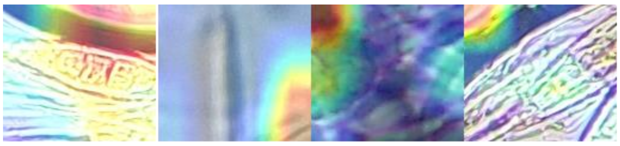


Figure 9. Examples of bad pattern recognition on blurry samples

V. Conclussion

In this paper, we described developing a identification model which, based on accuracy results and testing, presents a solid base for developing an application that can be used in practice and drastically accelerate the process of diagnostics.

Grad-CAM method used to visualize the decision making process has proven to be a very efficient method of evaluation of the model, not only in terms of validating the “thinking” process of the classifier, but to point out flaws and cases where errors happen.

Future development of the model and application will involve developing an algorithm that can reach the decision based on high resolution photo, from which number of smaller images will be cut, and then classification will be performed on each of the small sample images. This approach will increase precision of the diagnostics, since the decision will be a ruling of the mayor, rather than determination based on one small sample.

Acknowledgment

We would like to thank the Department of Microbiology and Immunology, Medical faculty, University of Niš, Serbia, for all resources, samples and advices given during the work on this project.

References

[1] Tasić S, Pešić S: Gljivičneinfekcije, dijagnoza I mogućnosti terapije, Punta Niš, Medicinski fakultet Niš, 2006 (odluka nastavno naučnog veća 14-376-5/2-2, 2006, udžbenik) R-12-5 M-44-2.

[2] Otašević S, Tasić-Miladinović N. Tasić A: Medicinska parazitologija-udžbenik sa CD-om. Univerzitet u Nišu- Medicinski fakultet Niš 2011 14-6735-6/2-4 2011.

[3] Valentina Arsić-Arsenijević, Marina Milenković, SuzanaOtašević, Dušan Pavlica. Medicinska mikologija i parazitologija. Društvo medicinskih mikologa Srbije, Centar za inovacije u mikologiji Beograd, V. Arsić-Arsenijević, Beograd 2012. ISBN 978-86-915391-1-5.

[4] Davise HL. Medically important fungi-a guide to identification, 3rd edn. Washington: American Society for Microbiology; 1995.

[5] S Tasić-Otašević , M Golubović, S Đenić, A Ignjatović, M Stalević, S Momčilović, M Bojanović , V Arsić-Arsenijević Species distribution patterns and epidemiological characteristics of otomycosis in Southeastern Serbia. J Mycol Med; 2020 Jun 30;101011. doi: 10.1016/j.mycmed.2020.101011.

[6] Otašević S, Momčilović S, Stojanović NM, Skvarč M, Rajković K, Arsić-Arsenijević V. Non-culture based assays for the detection of fungal pathogens. J Mycol Med. 2018 Jun;28(2):236-248. doi: 10.1016/j.mycmed.2018.03.001. Epub 2018 Mar 29.

[7] Pesic Z, Otasevic S, Mihailovic D, Petrovic S, Arsic-Arsenijevic V, Stojanov D, Petrovic M. Alternaria-Associated Fungus Ball of Orbit Nose and Paranasal Sinuses: Case Report of a Rare Clinical Entity. Mycopathologia. 2015 Aug;180(1-2):99-103. doi: 10.1007/s11046-015-9881-6. Epub 2015 Mar 7.

[8] Mina Milanović, Aleksandar Milosavljević, Determination of molds isolated from patient materials, based on the microscopic morphological characteristics, ICIST, March 2021 (proceedings in preparation)

[9] How to Load Large Datasets From Directories for Deep Learning in Keras by Jason Brownlee. Available online: https://machinelearningmastery.com/how-to-configure-image-data-augmentation-when-training-deep-learning-neural-networks/ (accessed on 20 May 2021).

[10] R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural Features for Image Classification,” IEEE Trans. Syst. Man. Cybern., vol. SMC-3, no. 6, pp. 610–621, Nov. 1973.

[11] J. A. Jose and C. S. Kumar, “Genus and Species-Level Classification of Wrasse Fishes Using Multidomain Features and Extreme Learning Machine Classifier,” Int. J. Pattern Recognit. Artif. Intell., Mar. 2020.

[12] P. J. D. Weeks, M. A. O’Neill, K. J. Gaston, and I. D. Gauld, “Species-identification of wasps using principal component associative memories,” Image Vis. Comput., vol. 17, no. 12, pp. 861–866, 1999.

[13] G. Csurka, C. R. Dance, L. Fan, J. Willamowski, and C. Bray, “Visual Categorization with Bags of Keypoints,” in Workshop on statistical learning in computer vision, ECCV, 2004.

[14] I. Kanellopoulos and G. G. Wilkinson, “Strategies and best practice for neural network image classification,” Int. J. Remote Sens., vol. 18, no. 4, pp. 711–725, Mar. 1997.

[15] Y. LeCun et al., “Handwritten digit recognition with a back-propagation network,” papers.nips.cc, pp. 396–404, 1990

[16] Y. LeCun and Y. Bengio, “Convolutional networks for images, speech, and time series.MIT Press, Cambridge,” Handb. brain theory neural networks, vol. 3361, no. 10, 1995

[17] EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, Mingxing Tan, Quoc V. Le, International Conference on Machine Learning, 2019

[18] Image_classification_efficientnet_fine_tuning. Available online:https://keras.io/examples/vision/image_classification_efficientnet_fine_tuning/ (accessed on 25 May 2021).

[19] EfficientNet: Improving Accuracy and Efficiency through AutoML and Model Scaling. Available online: https://ai.googleblog.com/2019/05/efficientnet-improving-accuracy-and.html(accessed on 25 May 2021).

[20] Python. Available online: https://www.python.org/ (accessed on 26 May 2021).

[21] Keras: The Python Deep Learning Library. Available online: https://keras.io (accessed on 26 May 2021).

[22] TensorFlow. Available online: https://www.tensorflow.org/ (accessed on 26 May 2021).

[23] Understanding RMSprop — faster neural network learning. Available online: https://towardsdatascience.com/understanding-rmsprop-faster-neural-network-learning-62e116fcf29a (accessed on 26 May 2021).

[24] Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, Dhruv Batra, Dec. 2019

Freelancing blockchain: A practical case-study of trust-driven applications development

Milan Radosavljevic, Aleksandar Pesic, Nenad Petrovic, Milorad Tosic

***Abstract*—Nowadays, a large amount of work is done by freelancers across various areas – from graphical design and music composition to data input and software development. However, many issues appear due to participation of several third parties together with different rules and policies imposed by different platforms. On the other side, the emerging blockchain technology provides the execution of transactions in a trustable, decentralized, but still transparent manner. In this paper, we demonstrate a case-study where blockchain is adopted to eliminate the barriers and make freelancing more convenient and profitable at the same time. As an outcome, a proof-of-concept implementation of blockchain-based freelancing platform relying on Ethereum and Solidity smart contracts is presented that provides practical pointers for trust0driven applications development.**

***Index Terms*— Blockchain, Ethereum, Solidity, Freelancers**

I. Introduction

In today's business world, everything is based on trust. Any monetary transaction, ownership or arrangement. This trust, however, is provided in a very specific way - by the role of a third party, i.e. an institution of trust. In money transactions these are banks, in ownership relations there are cadastres and similar state institutions, in the case of any type of contract, there are courts. The positive side of these institutions, i.e. third parties, is that all parties to all these agreements trust them and expect protection in case of any unexpected occurrences. On the other hand, the appearance of third parties brings with it a lot of negative effects, so you often end up in a waiting list in order to make payments or get a certificate of ownership of a real estate and the like. Mistakes made by these institutions themselves are also very common, and as a rule they fall on the common man as a burden. There is bureaucracy, inefficiency, mistakes, enormous costs that at some point completely make the role of these intermediaries meaningless. The question is, is it possible to exclude third parties from future business, and still preserve that positive factor that they brought with them. In the last few years, there has been a development of a technology called *blockchain*, whose primary idea is just this. Decentralized system with minimal costs, efficient, transparent and yet safe enough to take on the role of e.g. banks. [1]

Milan Radosavljevic is student at Faculty of Electrical Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia (e-mail: milan.radosavljevic@elfak.rs).

Aleksandar Pesic is student at Faculty of Electrical Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia (e-mail: alpesh@elfak.rs)

Nenad Petrovic is teaching assistant at Faculty of Electrical Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia (e-mail: nenad.petrovic@elfak.ni.ac.rs)

Milorad Tosic is full professor at Faculty of Electrical Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia (e-mail: milorad.tosic@elfak.ni.ac.rs)

The share of freelancers in today's business world is large. A large amount of work is completed by freelancers, and as a freelancer, everyone has the opportunity to work in a huge number of domains as an independent. So, people who are engaged in graphic design, writing stories, teaching foreign languages and, of course, programming, earn their living by freelancing. Current platforms for freelancers are safe and the most popular places where freelancers can find work are [2] [3] [4]. The problem that arises is the possession of a third party and some rules that must be followed on certain platforms. Blockchain provides an opportunity to eliminate these problems and as a new solution sets some new boundaries and presents some new problems that are obtained by introducing this solution.

In this paper, it is explored how the blockchain technology can be leveraged to eliminate barriers when it comes to freelancing. As main outcome of this research, we introduce prototype implementation of freelancing platform based on Ethereum blockchain technology and Solidity smart contracts.

II. Background

A. Blockchain

Blockchain is a "cryptographically secure transactional singleton machine with a shared state". Cryptographically secure means that the creation of digital currency is provided by complex mathematical algorithms that are practically impossible to break. With the help of these algorithms it is almost impossible to cheat the system (e.g. creating fake transactions, deleting transactions etc.). A transactional singleton machine means that there is one canonical instance of the machine responsible for all transactions created in the system. In other words, there is one global truth that everyone believes in. Shared status means that the status stored on this machine is shared and available to everyone.

B. Ethereum blockchain platform

Ethereum blockchain is practically a transaction-based state machine. As it is known, the state machine receives inputs and, based on the current state, passes into new states. With Ethereum's state machine, we start from the "initial state". This is practically the state before any transaction occurred on the network. When transactions are executed, the initial state passes to some final state. At any time, the final state represents the current state in the Ethereum network. The state of Ethereum has millions of transactions. These transactions are grouped into "blocks". Each block contains some set of transactions and each block is cryptographically chained together with the previous blocks, which can be seen in Figure 1.

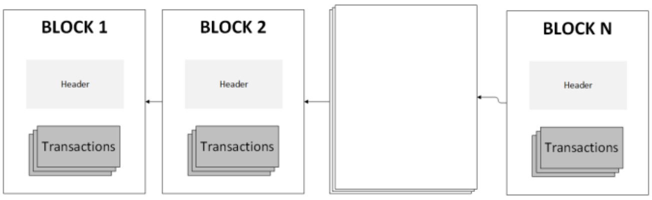


Fig. 1. Preview of Ethereum blockchain.

To cause a transition from one state to another, the transaction must be valid. For a transaction to be considered valid, it must go through a validation process known as mining. Mining is the process when a group of Ethereum nodes (more precisely computers) spend their computing resources to create a block of valid transactions. Any Ethereum network node that declares itself a miner can try to create and validate the block. All miners try to create and check blocks at the same time. Every miner provides mathematical "proof" when submitting a block, and this proof acts as a guarantee: if the proof exists, the block must be valid. The process of validation of each block by the miner who is supposed to provide mathematical proof is called *proof of work*.

C. Concept of fees

One very important concept in Ethereum is the concept of fees. Any calculation that occurs as a result of a transaction on the Ethereum network is charged a fee. The fee is paid in denominations called "gas". Gas is a unit used to measure the fees required for a particular calculation. Two factors determine how much it takes to pay for an action: the gas price, and how much gas that action requires. The important part is that Ethereum gas prices aren’t fixed. Gas prices are determined by supply and demand. The busier the Ethereum network, the higher the gas price. The amount of gas required for each transaction depends on how complex the transaction is. Gas prices are denoted in gwei, which itself is a denomination of ETH. Gwei is 10^{-9} ETH. It is possible to set a gas limit for each transaction. Gas limit refers to the maximum amount of gas you are willing to consume on a transaction.

D. Solidity smart contracts

A smart contract is a contract that is performed by itself together with the terms of the contract to which the parties have agreed. The terms of the contract are written directly in the code of that smart contract. And the contract itself and the 'consent' of the participants exists throughout the Ethereum network. Practically smart contracts are programs that are immutable and deterministic. They depend on the context of the Ethereum Virtual Machine and the decentralized global network. The contract controls the execution of transactions, which are public and non-refundable. In essence, contracts are reduced to programs, which are modeled on traditional contracts, ‘*if it happens, then do it*’. The contract is executed on many computers to ensure reliability and trust. Smart contracts provide autonomy, trust, speed, security and money savings.

III. Related Work

The blockchain was invented in 2008 by a group or an individual, and this information is still unknown to the public. Therefore, we can consider that blockchain, and at the same time Ethereum, is a newer technology that poses some new challenges and problems in front of us. Much research has been conducted in academia as well as in industry to explore the benefits of smart contracts as well as the worlds in which they are applicable. There are many smart contract platforms on the market with different features that suit certain applications. In [5], the authors focus on the technique of using blockchain to store vaccination records, which is secure and efficient and is based on smart contracts found on the Ethereum platform. In [6], a system based on blockchain was proposed, which refers to workers who are temporarily employed in companies. In that way, employees are provided with a fair and legal salary for their work obligations, as well as protection if the employer becomes a debtor. The author's work in [7] gives a focus on climate change and proposes a solution which, with the use of blockchain, would reduce global warming while keeping records of the crown impression of the product. In [8], the authors provided an overview of all the challenges that smart contracts would face in the future.

IV. Solution Overview

A. System Architecture

The application architecture consists of three parts: client side, Web3 interface and server side. The server side is located on the Ethereum blockchain network and uses Solidity smart contracts which are accessed via Web3 interface on the client side. The client side is located in the browser and uses HTML, CSS and Javascript programming language. Also, the client side contains the Web3.js Javascript library through which it communicates with Solidity smart contracts as can be seen in Figure 2.

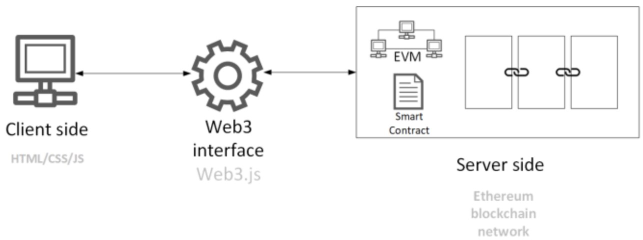


Fig. 2. Representation of system architecture.

B. Tools used for system implementation

1) Ganache

Ganache is a local Ethereum blockchain that runs on a local computer. Intended for the development and testing of smart contracts and decentralized applications in a secure and deterministic environment. Provides ten externally owned accounts for testing purposes. The application contains a graphical interface and can also be used as a console application.

2) Metamask

Metamask is a software that allows you to own a cryptocurrency wallet and allows you to interact with the Ethereum blockchain. Metamask provides the ability to store and manage account addresses on the browser. It also allows us to connect securely to decentralized applications. Using this software enables multi-user browser behavior.

3) Remix IDE

Remix IDE is an open source web and desktop application. It enables the rapid development of Solidity smart contracts and contains a large number of plugins as well as a graphical interface. The Remix IDE is used for contract development, but also as a platform for learning programming on the Ethereum platform. The Remix IDE is part of a Remix project that develops a handful of tools related to Solidity smart contracts. It is written in the Javascript programming language and allows you to run and test contracts in a web browser. It also allows you to test, debug and deploy contracts as well as many other useful options. [9]

V. Implementation

As indicated in the System Architecture chapter, the system consists of a server and a client side. The server side consists of smart contracts written in the Solidity programming language. The UML diagram in Figure 3 shows the organization of the contracts and the structures used in the system. The main component is a *FreelancerContract* contract that uses *Service*, *FreelancerStructure* and *Offer* data structures. This component represents one Freelancer who is registered in the system. While *PlatformContract* acts as a repository and it contains all registered freelancers in the system. In addition to the data structure that stores basic information about the freelancer, there are also structures for services and offers. *Service* is what a freelancer offers, while an *Offer* acts like a real job offer.

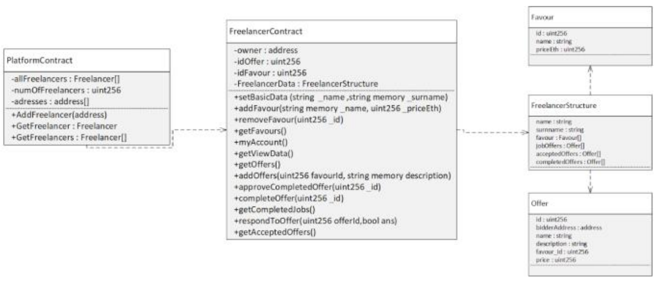


Fig. 3. UML contract diagram on Ethereum platform.

The client side of the application is written in the Javascript programming language. On the client side there are two proxy classes that correspond to the contracts on the server side, and also uses the Web3.js library to communicate with the server side. Practically one function on the client side calls one function in the contract. There are also functions that are handlers for the events which are broadcasted in contracts. The UML class diagram of the client side can be seen in Figure 4.

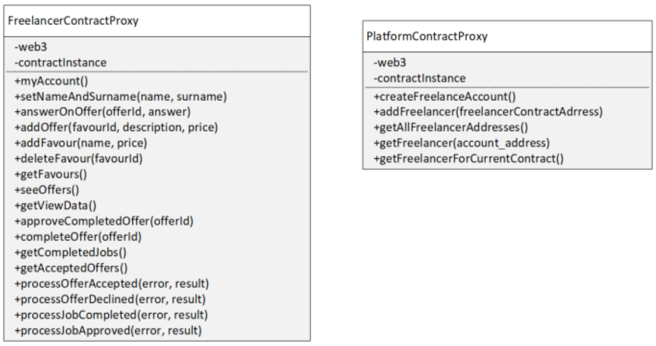


Fig. 4. UML class diagram on the client side.

VI. Results

Figure 5 shows the initial view of the client application. The application offers the ability to create a new account and navigation that allows to navigate through the entire application.

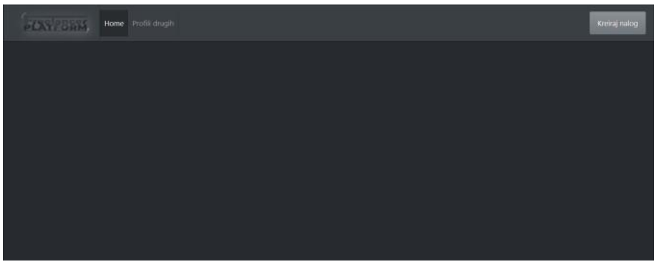


Fig. 5. Initial view of the application.

After creating the account, you get the view as in Figure 6 and the application allows the user to set the name and last name, add services he is offering and have an insight into the job offers that are offered to him, offers accepted by the user and offers waiting for client's approval.

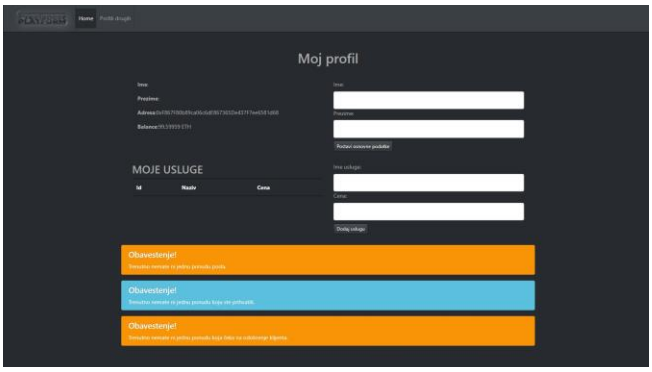


Fig. 6. View of the application after creating the account.

Clients can see the services of all freelancers on the Profiles of Others page, which can be selected from the navigation. The view of the page can be seen in Figure 7. The application provides the ability to search all offered freelancer services that are in the system and the ability to send a service offer to a specific freelancer that will be displayed on the homepage of the freelancer for whom the offer is intended.

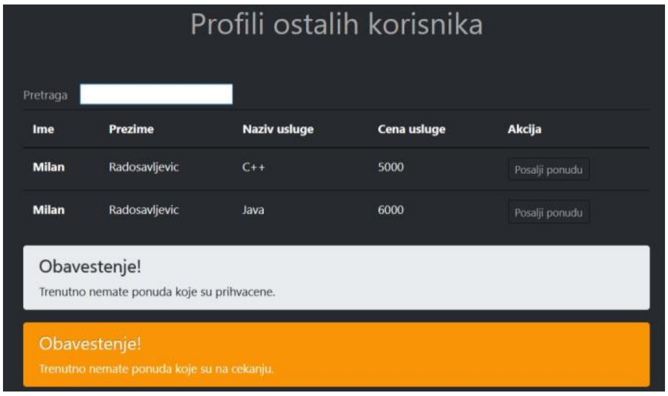


Fig. 7. View of page *Profiles of other users*.

When submitting an offer it is necessary to add a description and offer a price for a particular service. Based on the submitted offer, the freelancer will decide whether to accept the offer for the job. Form for sending the offer can be seen in Figure 8. After the offer is sent, the offered price is transferred from the client's account to the smart contract account. In case the freelancer rejects the offer, Ether will be returned to the customer account.

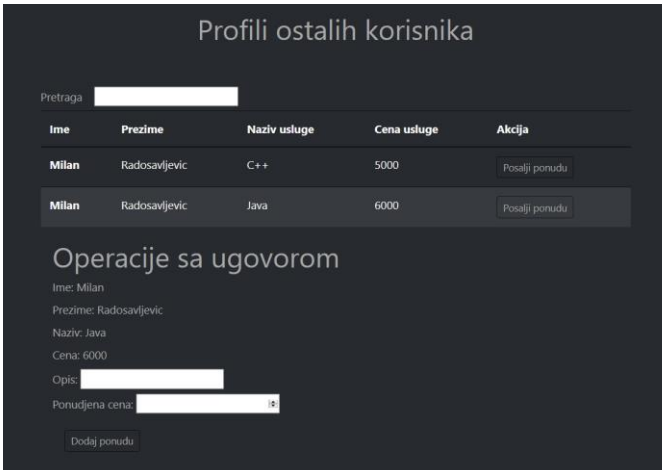


Fig. 8. Appearance of the offer submission form.

After the accepted offer, freelancer completes the offer and sends it to the client for approval, after the approved offer, the client accepts the completed work and only then the offered price is transferred to the freelancer's account. The layout of the offer table can be seen in Figure 9.

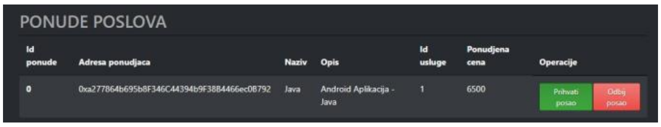


Fig. 9. Appearance of the table with job offers.

VII. Conclusions and Future Work

A platform for supporting a freelancing work community is developed using Solidity smart contracts on the Ethereum platform. The implementations process is used as a case-study for learning practical aspects of trust-driven applications development. Practical experiences and results are presented in this paper.

This paper gives our first experiences and more systematic approach is needed particularly validation and evaluation that are planned for future work. Regardless of absence of a full scientific rigor, we present our experiences that could be useful for future work in the field of trust-based applications development.

The good side of the solution presented in this paper is that it exploits advantages of the blockchain such as reliability, security and speed. Due to the fact that since its launch, the Ethereum platform has not had downtime due to consensus and decentralized approach. Hence the fact that this is another advantage, more precisely the advantage that the platform is always online and working. This further implies that the applications running on the blockchain do not have a downtime, including this one. The solution has an intuitive user interface that is capable of expansion. Even though the application was developed as a proof-of-concept, it shows high potential for commissioning in a real environment.

One of disadvantages of the approach is that every action that changes the state of the blockchain uses gas that requires compensation from the user, as stated in the chapter *Background*. Hence, some actions performed in the application are not free. In the current prototype, freelancer can not deliver product to the client who hired him. In future work this shortcoming could be fixed by connecting accounts with Github service, for example. It is possible to further optimize the speed of the application, on both client and server side, and to minimize the fee spent for performing actions on the server side. The obtained solution proves that it is possible to reach a satisfactory solution at an acceptable price.

Acknowledgment

This work has been supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

References

[1] Uvod u blockchain [online], Blog, Đorđe Ivanović, 2018. Source: https://blog.itkonekt.com/2018/07/30/uvod-u-blockchain/
[2] Freelancer [online], freelance recruitment platform, https://www.freelancer.com
[3] Upwork [online], freelance recruitment platform, https://www.upwork.com
[4] Fiverr, freelance recruitment platform [online], https://www.fiverr.com
[5] S. K. Deka, S. Goswami, A. Anand, "A Blockchain Based Technique for Storing Vaccination Records", IEEE Bombay Section Signature Conference (IBSSC), pp. 135-139, 2020.
[6] A. Pinna, S. Ibba, "A blockchain-based Decentralized System for proper handling of temporary Employment contracts", Proceedings of the 2018 Computing Conference vol. 2, pp. 1-6, 2019.
[7] A. M. Rosado da Cruz, F. Santos, P. Mendes, E. Ferreira Cruz, "Blockchain-based Traceability of Carbon Footprint" Proceedings of the 22nd International Conference on Enterprise Information Systems (ICEIS), pp. 1-10, 2020.
[8] Sh. Wang, Y. Yuan, X. Wang, J, Li, R. Qin, F. Wang, "An Overview of Smart Contract: Architecture, Applications, and Future Trends", IEEE Intelligent Vehicles Symposium, pp. 108-113, 2018.
[9] Remix IDE documentation [online]. Source: https://remix-ide.readthedocs.io/en/latest/

Comparative analysis of intra-board synchronous serial communication interfaces

Predrag Petronijević, *Vlatacom Institute of High Technologies, Milutina Milankovica 5, 11070 Belgrade, Serbia*
Vladimir Kuzmanović, *Faculty of Mathematics, University of Belgrade*

***Abstract* — Designing custom-made hardware for special purposes is a challenging process. During the development, it is essential to take into consideration the required performance of the device, component availability on the market as well as the final price of the developed and assembled product. Almost every modern hardware consists of various sensors, memories, AD/DA converters and a microcontroller to control and manage the interaction off all those devices. Based on the purpose of the device being developed, the engineer has to make a decision on the components that will be used in the final product. For this decision to be justifiable, the engineer needs to have a very high level of knowledge regarding the intricate world of interfaces required to establish the intercommunication of the components inside the device. Modern sensors, memories and AD/DA converters usually require some form of a high-speed serial interface, synchronous or asynchronous. In this paper we will analyze the three most commonly used serial synchronous communication interfaces: I^2C, SPI and SPORT. Also, we will explain the hardware and software properties and limits of every mentioned synchronous serial interface. Finally, the benefits and drawbacks of the chosen communication interfaces will be considered and conclusions drawn.**

***Index Terms* — computer engineering, embedded systems, sensors, synchronous serial communication**

I. Introduction

One aspect of designing new hardware is defining its application and the other aspect is defining a set of features the final product has to meet. The desired set of features can be divided into a set of operational and environmental limits, e.g. thermal resistance or voltage, and a set of desired performance characteristics, e.g. bandwidth or noise levels. This set of features limits the number of possible components that can be used in the design of the hardware. Even when limited with operational and performance characteristics, the choice of available hardware components is enormous due to a large number of manufacturers. Making the correct choice of hardware in order to meet the desired characteristics requires extensive knowledge [1-2]. One key decision to make is the choice of the right communication interface that will be used for intercommunication of the chosen components.

Predrag Petronijević is with Vlatacom Institute of High Technologies, Milutina Milankovica 5, 11070 Belgrade, Serbia (e-mail: predrag.petronijevic@vlatacom.com).

Vladimir Kuzmanović is with the Faculty of Mathematics, University of Belgrade, Studentski trg 16, 11000 Belgrade, Serbia (e-mail: vladimir_kuzmanovic@matf.bg.ac.rs).

Modern devices usually consist of various sensors, memories, AD/DA converters and many other components controlled by a microcontroller. This control is achieved by establishing intercommunication between the microcontroller and every component inside the device. In modern devices, this communication is digital and standardized to conform to one or more of the standard communication interfaces in use today [3-4].

Interfaces used today can be divided into categories based on the way the data is transferred between the devices. Two criteria can be used for this division. The first criterion is defined by the number of channels used in the transmission of the data. If the data is transmitted bit by bit in a specific order over a single channel, such transmission is called serial. If the data is sent as multiple bits at the same time over multiple channels, such transmission is called parallel. The other criterion is defined by the way the data is sent. If the data is sent in the form of a byte or a single character with start and stop bits added to the data, such transmission is called asynchronous because it does not require synchronization. If the data is sent in the form of groups or frames, such transmission is called synchronous because it requires synchronization between sender and receiver. Synchronous transmission is more reliable and full-duplex, while asynchronous transmission is half-duplex [5].

Inter-Integrated Circuit (I^2C) was discussed by Patel et al. [6], Lynch et al. [7] and Blum [8]. Wootton in [9] described the use of Serial Peripheral Interface (SPI) as a means of communication between the CPU and various peripheral devices. Gay in [10] described the properties of SPI and its operation was described by Dogan in [11]. SPI and I^2C were compared in [12-13]. SPI, I^2C and UART were analyzed in [14-15].

In this paper we will address and compare the three most commonly used synchronous serial communication interfaces for intercommunication between various devices. Besides the well-known and widely used I^2C and SPI protocols, we will also introduce Analog Devices proprietary SPORT protocol and perform comparative analyses of the three serial protocols. Section 2 of the paper introduces all three interfaces with their hardware and software properties and requirements. The next section analyses the benefits and drawbacks of these serial interfaces. Finally, section 4 draws conclusions on serial interfaces described in the paper.

II. Synchronous serial communication

A serial communication protocol in which data is sent as a continuous stream at a constant rate is described as synchronous serial communication. For communication to be called synchronous it is required that the clocks are synchronized in both the transmitting and receiving devices. The term synchronized refers to the clocks running at the same rate, which enables the receiver to sample the signal at the same intervals used by the transmitter. Synchronization of clocks permits the omission of start and stop bits. As a consequence, more information can be passed over a circuit per unit of time than with asynchronous serial communication.

Serial communication can be established via a communication channel or a computer bus. It is mostly used for long-distance communication and computer networks where parallel communication is impractical. The development of technology has made serial computer buses more common at shorter distances, mostly as a basis for cheap and simple intra-board communication between two or more integrated circuits on the same printed circuit board connected by signal traces and not external cables.

The three most commonly used synchronous serial communication protocols for intra-board communication are inter-integrated circuit (I^2C), serial peripheral interface (SPI) and Analog Devices synchronous serial peripheral port (SPORT).

A. Inter-Integrated Circuit (I^2C)

Inter-Integrated Circuit (I^2C) is a synchronous, multi-master, multi-slave, packet-switched and single-ended serial communication bus invented in 1982 by Philips Semiconductors. Today it is widely used for interfacing with lower speed peripheral integrated circuits from a microcontroller in short distance intra-board communication.

I^2C emphasizes design simplicity and low manufacturing costs over speed. It is usually used for accessing low-speed AD/DA converters, controlling small displays, reading diagnostic sensors, etc. I^2C enables the microcontroller to control a network of devices with just two general-purpose input-output pins and software. Many other serial protocols offer similar functionality but require more pins and signals to interconnect multiple devices [16].

Hardware requirements for establishing I^2C communication are rather simple. Two bidirectional open collector or open drain lines with typical voltages of +5V or +3.3V are required for connecting the devices. These two lines are called Serial Data Line (SDA) and Serial Clock Line (SCL). I^2C bus speed can range from 10 kbit/s to 5 Mbit/s depending on the revision of the protocol. The bit rate is defined for transfer between master and slave without taking into consideration any protocol overhead. The overhead includes a slave address and usually a register within the slave device and finally per byte acknowledge (ACK/NACK) bits. This makes the actual bitrate lower than the bitrate used would imply. High-speed I^2C is widely used in embedded systems, while lower speed version is used in personal computers.

The reference design is a bus with clock (SCL) and data (SDA) lines with 7-bit addressing to which the devices are connected. Devices connected to the bus are referred to as nodes. The number of nodes is limited by the address space and by the total bus capacitance of 400pF. This restricts communication distances to a few meters. In practice, I^2C is restricted to intra-board communication due to its relatively high impedance and low noise immunity which requires a common ground potential.

There exist two roles for the node on the bus: master and slave. The device is referred to as the master if it generates the clock and initiates communication with the slaves. The device is referred to as the slave if it receives the clock and responds when addressed by the master. The protocol supports multiple masters and multiple slaves on the same bus. Also, the roles of the device can be changed during its operation.

The protocol defines four modes of operation for a given device on the bus: master transmits, master receives, slave transmits and slave receives. Usually, each device on the bus will use a single role with two predefined modes of operation. Besides 0 and 1 data bits, the I^2C defines special signals which represent message delimiters. These signals are called START and STOP signals which are distinct from data bits. The communication between devices is as follows:

- The master is in master transmit mode and initiates the transmission by sending the START signal followed by a 7-bit address of the slave it wants to communicate with which is followed by a single bit designating whether the master wants to write to or to read from the slave.
- If the slave with the given address exists on the bus it responds with the ACK bit for that address. Then, the master continues to transmit either in transmit or receive mode according to the bit set while the slave continues in complementary mode.

The address and the data over the I^2C bus are sent in MSB mode. The START signal is a high-to-low transition of the data line (SDA) with the clock (SCL) line high. The stop signal is a low-to-high transition of SDA with SCL high. All other transitions of SDA take place with SCL low. The device which is in transmitting mode writes the data byte by byte to the SDA line. The device in receive mode sends the ACK bit after every byte. I^2C transmission may consist of multiple messages. The master terminates a message with a STOP signal if it is the end of the transaction. If the master wants to retain control of the bus for another message it sends another START signal.

I^2C physical layer is shown in Figure 1.

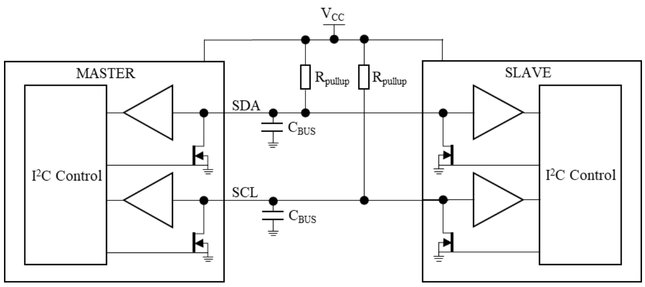


Figure 1. I^2C physical layer

B. Serial Peripheral Interface (SPI)

Serial Peripheral Interface (SPI) is a synchronous serial communication interface developed by Motorola in the mid-1980s [17]. It is used for short-distance communication in embedded systems. It is typically used for interfacing with memories, liquid crystal displays, sensors and AD/DA converters.

Devices communicating over SPI are organized as a master-slave architecture with a single master. The communication is achieved in full-duplex mode. The master device creates the frames for reading and writing. SPI supports multiple slave devices through selection with individual slave select lines. Sometimes, these lines are called chip select (CS) lines. The SPI bus specifies four logic signals:

- Serial clock (SCLK) – output from the master.
- Master Out Slave In (MOSI) – data output from the master.
- Master In Slave Out (MISO) – data output from the slave.
- Slave/Chip Select (SS/CS) – output from the master, active low.

For the communication to be established between devices, MOSI on a master device connects to MOSI on a slave device. Slave/Chip Select line is used instead of software addressing concept. Sometimes, MOSI on a slave device is labeled as Serial Data In (SDI) and MISO is labeled as Serial Data Out (SDO). This signal naming convention is used as an unambiguous way of labelling the pins of master and slave devices.

The SPI bus can operate with a single master device and one or more slave devices. Most slave devices have tri-state outputs so their MISO becomes high impedance when the device is not selected. This allows multiple slave devices to share common bus segments with each other.

For the communication to start, the master device has to configure the clock signal using a frequency supported by the slave. Then the master has to select the desired slave device with the logic level 0 on the appropriate SS/CS line. If the slave device requires a waiting period, the master device has to wait for at least that period of time before it starts issuing clock cycles on the SCLK line. During each cycle on the SCLK line, a full-duplex transmission occurs. The master sends a bit on the MOSI line and the slave reads it, while the slave sends a bit on the MISO line and the master reads it. This form of operation is maintained even when one-directional data transfer is intended.

Besides configuring the clock frequency, the master also needs to configure the clock polarity (CPOL) and clock phase (CPHA) with respect to the data. CPOL determines the clock polarity. CPOL value 0 defines a clock signal which idles at logic level 0 and each cycle consists of a pulse of 1. This translates to the leading edge being rising and the trailing edge is falling. CPOL value 1 defines the opposite. The clock idles at logic level 1 and each cycle consists of a pulse of 0. CPHA determines the timing of the data bits relative to the clock pulses. CPHA value 0 defines that the “out” side changes the data on the trailing edge of the preceding clock cycle, while the “in” side captures the data on the leading edge in the clock cycle. CPHA value 1 defines the opposite. The “out” side changes the data on the leading edge of the current clock cycle while the “in” side captures the data on the trailing edge of the clock cycle.

Finally, SPI supports word sizes that are not limited to 8-bit words but can range up to 32-bit words. Also, message size is arbitrary, as is its contents and purpose. The signal lines are shared between multiple devices, except for the slave select line which is unique per slave.

Its versatility, high speed and easy implementation coupled with board real estate savings compared to parallel buses have made it popular in many applications today. SPI interface is widely used in embedded systems for interfacing various sensors, control devices, memories and liquid crystal displays.

SPI physical layer is shown in Figure 2.

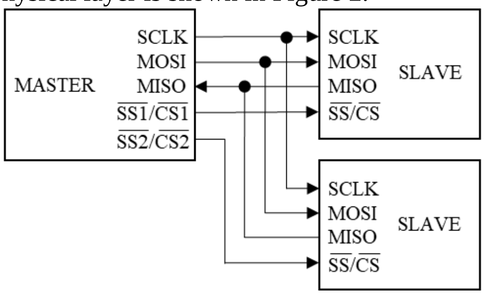


Figure 2. SPI physical layer

C. Synchronous Serial Peripheral Port (SPORT)

Synchronous Serial Peripheral Port (SPORT) is Analog Devices proprietary synchronous serial communication interface that supports a variety of serial data communication protocols. Key features of SPORT are continuously running clock and serial data words from 3 to 32 bits in length either most- or least-significant bit first. The protocol also supports two synchronous transmit and two synchronous receive data signals which double the total supported data stream. Finally, frames are synchronized with configurable synchronization signals [18].

For the SPORT interface to be established between two devices, the standard defines the following eight signals:

- Transmit Data Primary (DT0)
- Transmit Data Secondary (DT1)
- Transmit Clock (TSCLK)
- Transmit Frame Sync (TFS)
- Receive Data Primary (DR0)
- Receive Data Secondary (DR1)
- Receive Clock (RSCLK)
- Receive Frame Sync (RFS)

The values for clocks are independent and can be calculated by dividing the SCLK of the microcontroller with the correct value. The SPORT clocks are calculated with the following formula:

$$SPORTCLK = \frac{SCLK}{(2 \cdot (SPORTCLKDIV + 1))}$$

The smallest value the divisor SPORTCLKDIV can have is zero and the greatest value is 65535. TSCLK and RSCLK are independent and thus can have different values of SPORTCLKDIV. Depending on the value of SCLK and SPORTCLKDIV, the clock values for SPORT can be as high as 60 MHz or as low as 1 kHz. By default, the primary transmit and receive channels are enabled while the secondary transmit and receive channels are disabled.

Frame sync signal can be divided into early frame sync and late frame sync. Early frame sync is active for one clock pulse and then deactivates. Once the signal has been deactivated, valid data will be available. Late frame sync signal frames valid data and is active for the length of time that valid data is available. The signal is deactivated once the word to transmit or receive is fully sent.

SPORT protocol is proprietary and is supported by a majority of Analog Device microcontrollers and various types of integrated circuits for numerous applications. Such applications range from AD/DA converters, sensors, memories, health applications, smart industries, etc. Also, with a range of clock and frame synchronization options, the SPORT interface allows a variety of serial communication protocols and provides a glueless hardware interface to many industry-standard data converters and CODECs [19-20].

SPORT physical layer is shown in Figure 3.

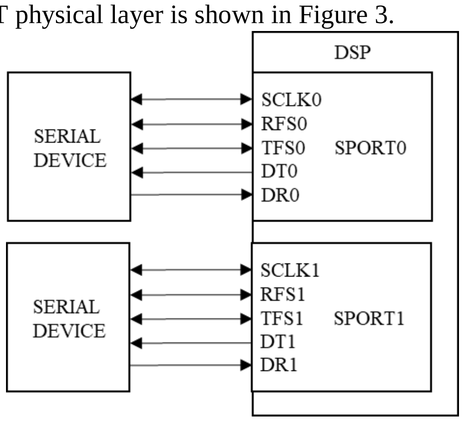


Figure 3. SPORT physical layer

III. Comparative Analysis

I^{2}C, SPI and SPORT all are synchronous bidirectional serial interfaces with considerable differences. The first obvious difference is the number of signals needed to establish communication between devices. The signals and number of lines required for establishing communication with each interface are displayed in table 1.

Table 1. Signals required for establishing communication

I^{2}C	SPI	SPORT
SDA Serial Data	MOSI Master Out Slave In	DT Serial Data Transmit
SCL Serial Clock	MISO Master In Slave Out	DR Serial Data Receive
	SCLK Serial clock	TFS Transmit Frame Sync
	SS Slave select	RFS Receive Frame Sync
		TCLK Transmit Clock
		RCLK Receive Clock

Considering the number of signals it is obvious that SPI and SPORT are full-duplex, while I^{2}C is half-duplex. Also, one other property to note is that I^{2}C is a multi-master multi-slave interface, while SPI and SPORT are single-master multi-slave interfaces.

Data transfer should also be considered when choosing the protocol to be used in the final product. The limits for data transfer are displayed in table 2.

Table 2. Data transfer limits

I^{2}C	SPI	SPORT
100 kbit/s – 5 Mbit/s Predefined values depending on version	Depending on the implementation Usually in range n x MHz to 10n x MHz n – number of devices connected to a single master	SCLK/2 Mbit/s SCLK – processor clock frequency

The advantages of I^{2}C over SPI and SPORT are the ease of linking multiple devices and the fact that cost and complexity do not scale up with the number of devices. The limitation of I^{2}C is numerous. The first is its slave addressing scheme and its relatively low number of possible addresses which may lead to address collisions. One other limitation is the number of supported speeds which need to conform to a certain standard. Since I^{2}C is a shared bus there exists a possibility that a single device could hang the entire bus. This happens if any device holds the SDA or SCL lines low, which prevents the master from sending START and STOP signals and reset the bus. Also, starvation is possible where a slower device starves the bandwidth needed by faster devices and thus increases latencies when other devices are addressed. Taking all this into consideration it is advisable to use I^{2}C for communication with on-board devices that are accessed only occasionally with no need for low latencies and high-speed bidirectional communication.

The advantages of SPI over I^{2}C and SPORT are complete protocol flexibility with variable size words and arbitrary choice of message size, contents and purpose. Also, hardware interfacing is easy. Slaves do not need a unique address since they are addressed with a per slave chip select line and slave devices do not need precision oscillators since they use the master's clock. Disadvantages compared to I^{2}C are the increased number of pins required for communication and the lack of slave ACK which enables the master to transmit data to nowhere without knowing it. Also, SPI protocol supports only one master, does not have a formal standard so validating conformance is impossible and does not support dynamically adding nodes. Taking all this into consideration, SPI is applicable in situations where the data transfer is organized in packets of arbitrary size and full-duplex. Also, it is applicable when there are a number of slaves communicating with the same SPI modes, because frequent changes of SPI mode severely impact the performance of communication.

Compared to the other two protocols, the main advantage of SPORT protocol is the support for multichannel transmits and receives of up to 128 channels. Also, a wide selection of data sizes is also a benefit as is the programmable polarity of both frame sync signals and data receive and transmit clocks. Finally, significantly higher data rates and double-buffered data registers that allow continuous data stream are a big advantage compared to both SPI and I^{2}C. The main disadvantages of SPORT are the fact that it is proprietary and supported only by Analog Devices products and that the complexity of supporting software components can be higher than that of competing schemes.

IV. Conclusion

In this paper we presented the three most commonly used synchronous serial protocols. The introduction showed that the engineer needs to have a broad knowledge regarding communication protocols to be able to make the right choice on the protocol to be used with respect to the desired operational and performance limits as well as to justify the proposed design. I^2C, SPI and SPORT are presented in detail and their properties, requirements and applications are discussed. Finally, the benefits and drawbacks of all three mentioned protocols are compared and analyzed which led to the conclusion on the suitability of the protocols in various scenarios. In the future, we intend to further research asynchronous communication protocols and their properties as well as inter-board communication protocols. We will focus on Controller Area Network (CAN) and Universal Asynchronous Receive Transmit (UART).

Acknowledgment

The research is founded by the Vlatacom Institute of High Technologies under project *#161 V155MM*.

References

[1] J. Staunstrup, W. Wolf, "Hardware/Software Co-Design: Principles and Practices," Springer Science & Business Media, 1997.

[2] P. Horowitz, W. Hill, "The Art of Electronics, 3rd edition," Cambridge University Press, 2015.

[3] J. Cowley, "Communications and Networking: An Introduction," Springer, 2007.

[4] IBM Corporation, "Data Communications Primer," Form C20-1668-0.

[5] J. Patrick, "Serial Protocols Compared," Embedded Staff, May 31, 2002, available at: https://www.embedded.com/serial-protocols-compared/.

[6] S. Patel, P. Talati, S. Gandhi, "Design of I2C Protocol," International Journal of technical innovation in Modern Engineering & Science, Vol 5, no. 3, pp. 741-744, 2019.

[7] K. M. Lynch, N. Marchuk, M. L. Elwin, "I^2C communication," Embedded Computing and Mechatronics with the PIC32, Newnes, 2016.

[8] J. Blum, "The I^2C Bus," Exploring Arduino; Tools and Techniques for Engineering Wizardry, 2nd Edition, 2019.

[9] C. Wootton, "Serial Peripheral Interface (SPI)," Samsung Artik Reference, Apress, Berkeley, 2016.

[10] W. W. Gay, "SPI Bus," Mastering the Raspberry Pi, Apress, Berkeley, 2014.

[11] I. Dogan, "Serial Peripheral Interface Bus Operation," SD Card Projects Using the PIC Microcontroller, Newnes, 2010.

[12] F. Leens, "An introduction to I^2C and SPI protocols," IEEE Instrumentation & Measurement Magazine, vol. 12, no.1, pp. 8-13, 2009.

[13] D. V. Gadre, S. Gupta, "Serial Communication: SPI and I2C," Getting Started with Tiva ARM Cortex M4 Microcontroller, Springer, 2017.

[14] A. Subero, "USART, SPI and I2C: Serial Communication Protocols," Programming PIC Microcontrollers with XC8, Apress, Berkeley, 2017.

[15] S. Shanthipriya, S. Lakshmi, "Design and verification of low speed peripheral subsystem supporting protocols like SPI, I2C and UART," ARPN Journal of Engineering and Applied Sciences, vol. 12, pp. 7368-7391, 2017.

[16] A.K. Oudjida, M.L. Berrandjia, R. Tiar, A. Liacha, K. Tahraoui, "FPGA Implementation of I2C & SPI Protocols: a Comparative Study," DOI: 10.1109/ICECS.2009.5410881, 2009 16th IEEE International Conference on Electronics, Circuits and Systems - (ICECS 2009), Yasmine Hammamet, Tunisia, 13-16. December, 2009.

[17] Motorola, Freescale, NXP, "SPI Block Guide v3.06," 2003.

[18] B. Prabhalika, M. Kiran Kumar, "Fpga implementation of Design and verification Synchronous serial port(S-PORT)," ISSN: 2321-9939, International Journal of Engineering Development and Research IJEDR, India, 2013.

[19] A. Vasudev Prabhugaonkar, J. Rayala, "Interfacing AD7676 ADCs to ADSP-21365 SHARC® Processors," Engineer-to-Engineer Note EE-248, Analog Devices, Rev 1, October 7, 2004.

[20] "Implementing UART Using the ADuCM3027/ADuCM3029 Serial Ports," Application Note AN-1435, Analog Devices, Norwood, MA, 2017.

РОБОТИКА И ФЛЕКСИБИЛНА АУТОМАТИЗАЦИЈА
/
ROBOTICS AND FLEXIBLE AUTOMATION

(PO/ROI)

Elektrotaktilni *feedback* za prepoznavanje osobina predmeta manipulisanih mekim robotom

Gorana Marković, Jovana Malešević, Milica Isaković, Miloš Kostić, Matija Štrbac, Kosta Jovanović

***Apstrakt*— Glavni preduslov da se u potpunosti iskoriste prednosti telemanipulacionih sistema jeste dvosmjerna razmjena haptičkih informacija između operatora i udaljene okoline, što omogućuje operatoru da percepira kolizije, kontaktne sile, težinu, oblik, veličinu, strukturu objekta i sl. Da bismo odgovorili na neke od ovih zahtjeva, u ovom radu predlažemo upotrebu elektrotaktilne stimulacije s prostornim i frekvencijskim kodiranjem informacija. Korišćeni sistem sastoji se od robotskog aktuatora konačne mehaničke krutosti, električnog stimulatora i površinske elektrode s više polja (engl. *multi-pad*) koja se postavlja na vrh kažiprsta subjekta. Za razliku od sličnih postavki haptične povratne sprege, u našoj studiji ispitanici su se oslanjali isključivo na taktilne, bez vizuelnih ili auditivnih povratnih informacija. Eksperimentalni rezultati pokazali su da elektrotaktilna stimulacija može poslužiti za prenos informacija o mekoći (prosječna stope prepoznavanja 3 nivoa mekoće iznosila je 90%) i veličini (prosječna stopa prepoznavanja 2 veličine iznosila je 98%) predmeta koji se hvata udaljenim aktuatorom.**

***Ključne reči*—Telemanipulacija; Elektrotaktilni feedback; Meki robot**

I. UVOD

Daljinski upravljani (telemanipulacioni, teleoperacioni) robotski sistemi koriste se u sredinama u kojima direktni ljudski kontakt može biti ili nebezbjedan (poput nuklearnih postrojenja [1] i svemira [2]) ili nedovoljno precizan (robotski asistirana telehirurgija [3]). Kontrola ranih telemanipulacionih sistema bazirala se prvenstveno na vizuelnim povratnim informacijama (eng. *feedback*), pri čemu ljudski operater nije dobijao povratne informacije o interakciji aktuatora sa okolinom, što je obično rezultiralo velikim kognitivnim opterećenjem za njega.

Da bi se nadoknadila ograničenja vizuelnih i auditivnih povratnih informacija, počevši od sredine 1990-ih, razvoj haptičkih interfejsa započet je džojsticima u industriji igara [4], a nastavio je brzi rast, posebno u smjeru vibrotaktilnog i *feedback*-a sile. Studije o učinku haptičkog *feedback*-a na zadacima manipulacije objektima otkrile su da haptička povratna sprega poboljšava performanse i efikasnost telemanipulacije utičući na smanjenje kontaktnih sila, potrošnje energije, vremena i broja grešaka prilikom izvršenja zadatka [3], [5]–[7]. Međutim, vizuelni senzori (najčešće različite vrste kamera) u udaljenoj okolini i dalje ostaju nezamjenjiv dio *feedback* sistema [8].

Prevalencija vibrotaktilne stimulacije u odnosu na druge modalitete *feedback*-a postoji i dalje u većini današnjih primjena, uprkos određenim nedostacima. U posljednjih 20 godina realizovani su vibrotaktilni *feedback* sistemi namijenjeni različitim primjenama bilo samostalno [9] ili u kombinaciji s drugim *feedback* modalitetima [10]. Za razliku od vibrotaktilnih, sistemi elektrotaktilnog *feedback*-a su efikasni u pogledu potrošnje energije, jednostavni su za proizvodnju i sposobni su proizvesti senzaciju čiji se parametri mogu pouzdano kontrolisati [11]. Ideja da se povratna informacija obezbjedi pomoću sistema koji putem površinskih elektroda stimuliše kožu strujama malog intenziteta, pojavila se prije više od 40 godina [11].

Takva ideja korištena je u *feedback* sistemima za proteze gornjih udova [12]–[15], za pojačavanje interakcije za korisnike virtuelne realnosti [16], te pri upravljanju robotskom šakom tokom zadataka poput izbjegavanja prepreka i klin-u-rupu (engl. *peg-in-a-hole*) [17]. Pamungkas i Ward su u svoja tri istraživanja kombinovali elektrotaktilnu stimulaciju sa sistemom za stereo viziju [18], [19] i infra crvenim (engl. *infra-red, IR*) senzorima daljine i senzorima sile postavljenim na hvataljku robota [17]. Identifikaciju predmeta pomoću elektrotaktilnog *feedback*-a ispitivalo je nekoliko istraživačkih grupa. Li i saradnici su testirali stopu tačnosti prepoznavanja tri objekta različite težine: lakog, srednje teškog i teškog, koristeći različite šeme kodiranja, uključujući amplitudsku, frekvencijsku i modulaciju širine impulsa [20]. Arakeri je sa svojom grupom pokazao da su ispitanici s povezom preko očiju uspješno naučili da primjene elektrotaktilne senzacije za razlikovanje 27 predmeta različite težine, širine i popustljivosti [21]. Chai i njegov tim su pokazali da ispitanici mogu prepoznati predmete u tri veličine, tri mekoće kao i četiri nivoa sile hvatanja tih predmeta sa relativno visokom stopom tačnosti koristeći prostorno, amplitudsko i mješovito kodiranje [22].

U ovom radu, pokazali smo da elektrotaktilna stimulacija može na intuitivan način pružiti informacije ljudskom operatoru o mekoći i veličini predmeta kog hvata udaljeni aktuator. Ispitanici su se oslanjali isključivo na taktilne senzacije, dok je dotok vizuelnih i auditivnih informacija bio onemogućen. Hvatanje je izvedeno pomoću aktuatora promjenive krutost (engl. *Variable Stiffness Actuator*), QBmove Maker Pro [23], a podaci o položaju i momentu sile hvataljke prikupljenih sa njenih senzora su kodirani lokacijom i učestalošću elektrotaktilne stimuacije na vrhu kažiprsta.

U prikazanoj studiji testirali smo sljedeće hipoteze:

1. Elektrotaktilna stimulacija može proizvesti dovoljan broj različitih taktilnih senzacija tako da pouzdano prenose

Gorana Marković, Kosta Jovanović – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar Kralja Aleksandra 73, 11020 Beograd, Srbija (e-mail: kostaj@etf.bg.ac.rs)

Gorana Marković, Jovana Malešević, Milica Isaković, Miloš Kostić, Matija Štrbac – Tecnalia Serbia Ltd, Deligradska 9/39, 11000 Beograd, Srbija (e-mails:{gorana.markovic} {jovana.malesevic} {milica.isakovic} {milos.kostic} {matija.strbac}@tecnalia.com).

informacije o dvije veličine i tri mekoće predmeta;

2. Prostorna i frekvencijska modulacija su nezavisne, a mješovito prostorno-frekvencijsko kodiranje je intuitivno i može se razlikovati bez međusobnog maskiranja;

3. Visok nivo tačnosti prepoznavanja (iznad 90%) može se postići bez dodatnih vizuelnih ili auditivnih povratnih informacija.

II. METOD

A. Postavka sistema

Postavku sistema čine:

1. QBmove maker pro aktuator i hvataljka;
2. Tactility stimulator;
3. Posebno dizajnirana Tactility *multi-pad* elektroda;
4. Personalni laptop računar (Intel Core i7-10510U CPU @ 1.8 GHz, 16 GB RAM) i monitor (dijagonale jednake 60.45 cm).

1) Qbmove maker pro aktuator

QBmove maker pro je bidirekcioni, antagonistički aktuator promjenjive krutosti, čija ekvilibrijumska pozicija izlazne osovine zavisi od pozicija dva motora *q1* i *q2* (1). Položaj izlazne osovine (x) direktno se mjeri enkoderom, dok se moment sile (τ) procjenjuje iz (2).

$$x_{eq} = \frac{q_1 + q_2}{2}. \tag{1}$$

$$\tau = k_1 * \sinh(a_1 * (x - q_1)) + k_2 * \sinh(a_2 * (x - q_2)). \tag{2}$$

Parametri k_1 = 0,0227 Nm, k_2 = 0,0216, a_1 = 6,7328 1/rad i a_2 = 6,9602 1/rad, baš kao i jednačine, dobijeni su iz tehničkog lista [24].

Aktuator komunicira s računarom putem USB komunikacije. Softver s jednostavnim grafičkim korisničkim interfejsom (engl. *Graphical User Interface, GUI*) za skeniranje portova, povezivanje, aktiviranje i podešavanje aktuatora preuzet je s veb stranice QB robotics [23]. Hvataljka korištena u ovoj studiji je štampana 3D štampačem i sastoji se iz dva dijela: nepokretnog i pokretnog. Pokretni dio je povezan direktno na osovinu, tako da je apertura hvataljke direktno proporcionalna položaju izlazne osovine.

2) Sistem za elektrotaktilnu stimulaciju

Elektrotaktilni sistem se sastojao od stimulatora i posebno dizajnirane *multi-pad* elektrode za površinsku stimulaciju. Razvojni Tactility stimulator (Tecnalia R&I, San Sebastijan, Španija) je strujno kontrolisan stimulator koji generiše simetrične, bifazne impulse čija je impulsna širina u opsegu od 30 μs do 500 μs (sa korakom od 10 μs), frekvencija u opsegu između 1 Hz i 200 Hz (sa korakom od 1 Hz) i amplituda u rasponu od 0,1 mA - 9 mA (sa korakom od 0,1 mA). Elektroda je proizvedena procesom sitoštampe komercijalnim mastilima za biomedicinsku upotrebu na komercijalnoj PET (polietilen tereftalat) podlozi. Provodni sloj napravljen je od srebra (srebro-srebro-hlorid, Ag/AgCl, proizvođača Henkel Electrodag 6037E SS), sa izolacionim premazom (Henkel Electrodag PF-455B) koji pokriva provodna polja.

Čitav sistem elektrotaktilne stimulacije, zajedno s QBmove aktuatorom i predmetima koji se koriste za hvatanje, prikazan je na Sl. 1. Elektroda i način njenog pozicioniranja na kažiprstu prikazani su na Sl. 2.

B. Kodiranje feedback-a

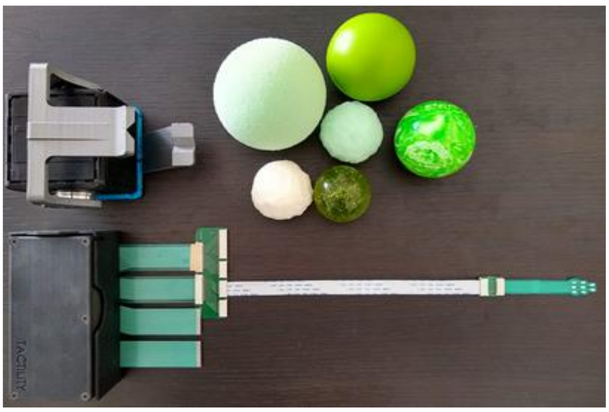
Informacija o mekoći predmeta je prenijeta putem frekvencijske modulacije stimulacije, dok su veličine kodirane prostorno. Prostorno kodiranje odnosi se na upotrebu različitih polja elektrode za izazivanje senzacija na različitim lokacijama, u ovom slučaju na vrha kažiprsta. Za ovaj eksperiment korištena su dva od osam polja elektrode: br. 2 i br. 7 (Sl. 2, lijevi panel). Pored toga, tri nivoa frekvencije su empirijski odabrana: niska = 5 Hz, srednja = 20 Hz i visoka = 50 Hz.

Prilikom prepoznavanja mekoće, maksimalni moment sile postignut tokom hvatanja procjenjen je pomoću (2) i zatim mapiran u jedan od tri nivoa frekvencije tako da najmanja mekoća odgovara najmanjoj frekvenciji. S druge strane, u zavisnosti od izmjerenog položaja izlazne osovine pokretnog dijela hvataljke u trenutku kontakta s objektom pali se određeno polje: br. 2 ako je predmet mali, br. 7 ako je predmet veliki. Granice nivoa momenta sile i pozicije određene su na osnovu pilot mjerenja prilikom hvatanja šest različitih lopti za unaprijed definisanu vrijednost krutosti aktuatora.

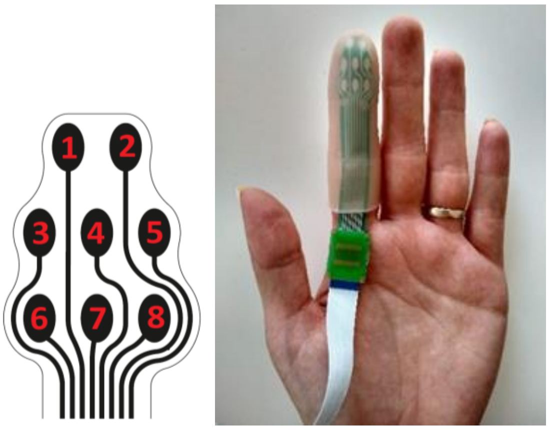
Tabela I rezimira prostornu i frekvencijsku mapu kodiranja.

C. Protokol

Deset zdravih dobrovoljaca (26,1 ± 3,7 godina, pol: 6 žena / 4 muškaraca, dominantna ruka: desna 9/lijeva 1) učestvovalo je u tri psihometrijska testa. Ispitanicima je objašnjen protokol i potencijalni rizici, nakon čega su potpisali pristanak za učešće u studiji. Ispitanici su bili smješteni u sjedeći položaj, s podlakticom oslonjenom na ravnu površinu i leđima okrenutim prema osobi koja vrši eksperiment, a samim tim i postavci sistema. Eksperiment je izveden u tihom ambijentu, a ispitanici su nosili izolacione slušalice u slučaju jače buke.



Sl. 1. Stimulator, *multi-pad* elektroda, Qbmove aktuator i objekti korišćeni u eksperimentu. Lopte su prečnika 5,4 cm (velika) i 4 cm (mala).



Sl. 2. Tehnički crtež elektrode (lijevo) i njeno pozicioniranje na vrh kažiprsta (desno).

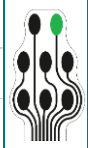
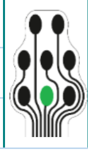
Osoba koja vrši eksperiment je postavila *multi-pad* elektrodu na kažiprst nedominantne ruke ispitanika i navukla silikonski naprstak preko elektrode (Sl. 2, desno). Silikonski naprstak je korišćen za obezbjeđivanje dobrog kontakta elektrode sa kožom i sprječavanje neželjenog pomjeranja elektrode. Prije tri glavna testa, protokol je uključivao kalibraciju i trening.

Kalibracija je podrazumjevala postupak ručnog podešavanja amplitude stimulusa kako bi se ispod oba polja izazvale jasne, ugodne i lokalizovane senzacije. Podešavanjem jednakih intenziteta senzacije ispod oba korišćena polja izbjegnuto je maskiranje slabijih stimulusa onim jačim. Preostali parametri stimulacije, frekvencija i širina impulsa, bili su konstantni tokom kalibracije intenziteta i postavljeni na 20 Hz, odnosno 400 µs, redom.

Trening se sastojao od familijarizacije i učenja sa podsticanjem. Tokom familijarizacije, u posebno pripremljenoj aplikaciji, ispitanik samostalno bira različite kombinacije parametara stimulacije (Tabela I), na taj način se upoznaje sa različitim osjećajima i vježba lokalizaciju osjećaja izazvanih aktiviranjem dva polja i diskriminaciju između tri nivoa frekvencije. Na osnovu samoprocjene ispitanika da je spreman, prelazi se na učenje sa podsticanjem. Učenje sa podsticanjem uključivalo je provjeru identifikacije šest poruka (2 lokacije i 3 frekvencije), dok je ispitanik primao usmene povratne informacije o tačnom odgovoru. Nakon deset uzastopno tačno identifikovanih poruka, ispitanik je smatran spremnim za testove. Trajanje treninga mjereno je za sve ispitanike.

Protokol testova obuhvatao je nekoliko koraka: osoba koja vrši eksperiment prvo stavlja 1 od 6 lopti (Sl. 1) između prstiju hvataljke, a potom pokreće robota putem specijalno dizajniranog LabVIEW (National Instruments, Austin, TX) grafičkog korisničkog interfejsa. Hvataljka se konstantnom brzinom zatvara dok ne dostigne unaprijed definisani položaj i uhvati predmet, a ispitanik prima taktilni *feedback* o veličini i/ili mekoći lopte i daje verbalni odgovor o svojoj interpretaciji primljenih povratnih informacija. Prvi i drugi test obuhvatali su po 10 hvatanja pseudo-nasumično odabranih loptica, dok je zadatak ispitanika bio da dâ usmeni odgovor o mekoći odnosno veličini lopte koju je robot

TABELA I
MAPA KODIRANJA SENZORSKIH PODATAKA U PARAMETRE ELEKTROTAKTILNE STIMULACIJE

Lopta	**Frekvencija [Hz]**	**Moment sile (Nm)**	**Aktivno polje**	**Pozicija izlazne osovine (°)**
Mala meka	5	$\tau < 1.6$		$x > 8$
Mala srednje meka	20	$1.6 < \tau < 2.2$		$x > 8$
Mala tvrda	50	$\tau > 2.2$		$x > 8$
Velika meka	5	$\tau < 1.6$		$x < 8$
Velika srednje meka	20	$1.6 < \tau < 2.2$		$x < 8$
Velika tvrda	50	$\tau > 2.2$		$x < 8$

uhvatio. Tokom prvog testa, polje br. 7 uvijek je bilo aktivno nezavisno od veličine predmeta, dok se frekvencija stimulacije mijenjala zavisno od mekoće lopte. Frekvencija stimulacije u drugom testu bila je 20 Hz, nezavisno od mekoće predmeta, dok se aktivno polje mijenjalo zavisno od njegove veličine. Treći test bio je kombinacija prethodna dva, i cilj mu je bio da ispita sposobnost istovremenog prepoznavanja i veličine i mekoće lopte u 12 pseudo-nasumičnih pokušaja. Ostale vrijednosti parametara (amplituda i širina impulsa) ostale su nepromijenjene u sva tri testa.

Nakon završetka eksperimenta, svaki od 10 ispitanika ocijenio je sljedeće tvrdnje na skali od 1 do 10 (1 - uopšte se ne slažem, 10 - u potpunosti se slažem):

Tvrdnja 1: Vjerujem da sam bio dobar u prepoznavanju mekoće (test 1)

Tvrdnja 2: Vjerujem da sam bio dobar u prepoznavanju veličina (test 2);

Tvrdnja 3: Vjerujem da je trening značajno doprinio uspješnosti u testovima.

III. ANALIZA PODATAKA

Glavna izlazna mjera u svim testovima bila je stopa uspješnosti prepoznavanja (SUP) (%) mekoće (test 1), veličine (test 2) ili obje osobine lopte (test 3). Pored toga, u testu 3 izračunali smo marginalne SUP za dvije osobine. S obzirom na to da je Anderson-Darling test pokazao da podaci nisu normalno raspoređeni, za statističku analizu su korišćeni neparametarski testovi, a rezultati u tekstu su prikazani u formi medijan (interkvartilni opseg - IKO). Wilcoxon-ov test rangova korišćen je za upoređivanje uspješnosti ispitanika prilikom prepoznavanja informacija o osobinama predmeta koje su prenijete pojedinačno i u kombinaciji, odnosno mekoće u testu 1 u odnosu na test 3 i veličine u testu 2 u odnosu na test 3. Prag statističke značajnost je postavljen na $p < 0.05$. Uticaj trajanja treninga na performanse testa ispitan je računanjem korelacije između ove dvije varijable.

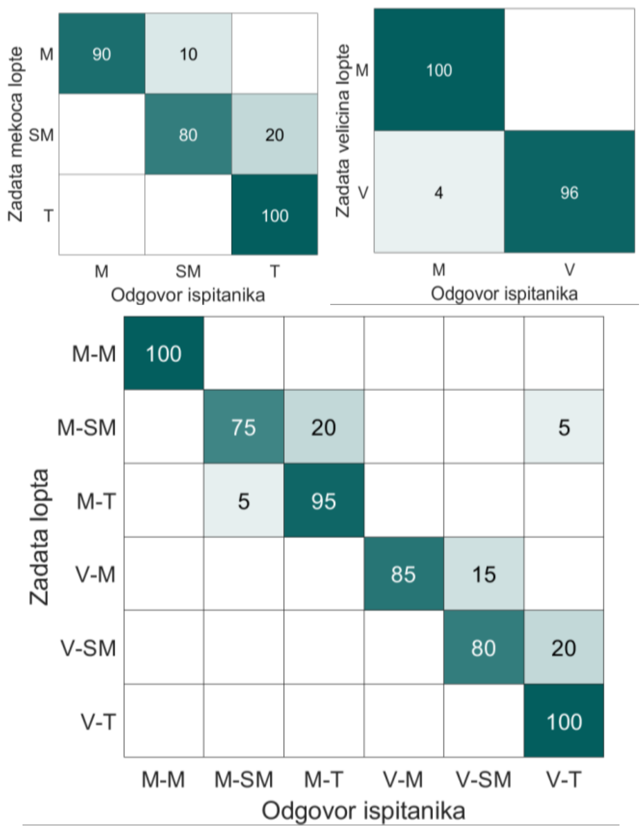
IV. REZULTATI

Sl. 3 prikazuje matrice konfuzije koje slikovito prikazuju rezultate ispitanika prilikom prepoznavanja mekoće lopte (test

1 - gore lijevo), veličine (test 2 - gore desno) i obje osobine u kombinaciji (test 3 - dole). U testu 1, svi ispitanici su svaki put prepoznali tvrdu loptu, dok su rijetko mijenjali srednje meku i meku loptu za onu jedan nivo tvrđu. U testu 2, svi ispitanici su prepoznali malu loptu u svim pokušajima, a jedini izvor greške bilo je pogrešno tumačenje velike lopte kao male u par navrata. U testu 3, ispitanici su češće pogrešno interpretirali mekoću lopte nego njenu veličinu, dok je SUP od 100% evidentna za najrazličitije lopte, tj. veliku tvrdu i malu meku.

Tabela II prikazuje pojedinačne i grupne rezultate po ispitanicima: SUP u tri testa, marginalne SUP prilikom identifikacije mekoće i veličine u testu 3, trajanje treninga i informacije o prethodnom iskustvu ispitanika u električnoj stimulaciji. Ispitanici su mogli prepoznati mekoću u testu 1 i veličinu u testu 2 sa vrijednošću medijane (IKO) SUP jednakom 100 (20)%, odnosno 100 (0)%. Marginalne SUP ove dvije osobine bile su vrlo slične SUP u pojedinačnim testovima: mekoća 96 (25)% i veličina 100 (0)%. Statistička analiza nije pokazala značajne razlike između SUP osobina u pojedinačnim i kombinovanom testu. Prosječno trajanje treninga bilo je ispod 5 minuta i pokazuje umjerenu negativnu korelaciju (*c* = - 0,49) sa SUP u testu 3.

Tabela III pokazuje rezultate upitnika. Subjektivno mišljenje o prepoznavanju veličine (Tvrdnja 2) ima veću



Sl. 3. Matrice konfuzije sa SUP (%) mekoće (gore, lijevo), veličine (gore, desno) i obje osobine zajedno (dole). Mekoća: ‘M’ = meka, ‘SM’ = srednje meka, ‘T’ = tvrda; Veličina: ‘M’ = mala, ‘V’ = velika; Kombinovano: ‘M-M’ = mala meka, ‘M-SM’ = mala srednje meka, ‘M-T’ = mala tvrda, ‘V-M’ = velika meka, ‘V-SM’ = velika srednje meka, ‘V-T’ = velika tvrda.

ocjenu od subjektivnog mišljenja o prepoznavanju mekoće (Tvrdnja 1).

V. Diskusija

A. Rezultati testova

Tabela III sugeriše da su subjekti bili uspješniji u prepoznavanju veličine, što odgovara prepoznavanju 2 lokacije stimulacije, u odnosu na prepoznavanje mekoće, što odgovara prepoznavanju 3 frekvencije stimulacije. To se može primjetiti upoređivanjem rezultata prvog i drugog testa, ali i na osnovu rezultata trećeg testa u kojem su ispitanici rijetko pogrešno protumačili informaciju o veličini objekta. Marginalna SUP mekoće i marginalna SUP veličine u kombinovanom zadatku u potpunosti su usklađene sa SUP postignutim u pojedinačnim testovima. Statistička analiza pokazala je da ne postoji značajna razlika u SUP u pojedinačnom i kombinovanom zadatku, čime je potvrđena početna hipoteza da se prostorna i frekvencijska modulacija mogu koristiti kao dva nezavisna *feedback* kanala. Prosječna SUP veličine (test 2) od 100 (0)%, bio je očekivan ishod s obzirom da je u prethodnim istraživanjima zabilježena prosječna tačnost od 97% za prostornu diskriminaciju 4 polja [12]. Međutim, budući da su u pomenutom istraživanju ova 4 polja bila locirana oko podlaktice s razmakom između polja jednakim ~ 4,8 cm, a u našoj studiji polja su smještena na vrhovima prstiju s razmakom od samo ~1,5 cm, ovi rezultati ohrabruju i sugerišu da bi i više lokacija stimulacije moglo pouzdano da se prepozna na vrhu prsta.

U zadatku prepoznavanja mekoće, ispitanici su takođe postigli visoku SUP od 100 (20)%. Ovaj rezultat je u skladu s prethodnim rezultatima [12] gdje su autori pokazali da se 3 nivoa frekvencije (4 Hz, 27 Hz i 100 Hz) mogu prepoznati sa SUP jednakom 99% nakon kratkog treninga. Na gornjem, lijevom panelu na Sl. 3 može se primjetiti da je najniža SUP od 80% postignuta prilikom prepoznavanja srednje meke lopte, koja je pogrešno prepoznata kao tvrda, dok su s druge strane svi ispitanici tačno prepoznali tvrdu loptu. U kombinovanom testu, SUP je bila najniža prilikom identifikacije srednje meke lopte, i velike (80%) i male (75%). Matrica konfuzije (Sl. 3, dole) pokazuje da su ispitanici često mijenjali srednje meku loptu za tvrdu. Štaviše, velika, meka lopta je relativno često interpretirana kao velika, srednje meka lopta. Generalno se više grešaka javljalo kada su se u testu koristile velike lopte.

B. Individualni rezultati i uticaj dužine trajanja treninga

Četiri od deset ispitanika imalo je SUP jednaku 100% u sva tri testa, a prosječnu SUP jednaku ili veću od 90% postiglo je 7 ispitanika u testu 1, 8 ispitanika u testu 2 i 6 ispitanika u testu 3, koji se smatra najtežim. Primjetno je da je svaki ispitanik koji je griješio u prvom testu, gotovo sigurno postigao nižu SUP u trećem testu, što opet potvrđuje da je prepoznavanje mekoće bio izazovniji zadatak. Uzimajući u obzir subjektivni utisak, može se primjetiti da su ispitanici mogli dobro procjeniti sopstvene performanse u testovima (Tvrdnje 1 i 2, Tabela III). Na primjer, ISP8 je ocjenio

TABELA II

SUP (%) U TRI TESTA, MARGINALNA SUP U TESTU 3, TRAJANJE TRENINGA I INFORMACIJE O PRETHODNOM ISKUSTVU 10 ISPITANIKA. USREDNJENE VRIJEDNOSTI SU PRIKAZANE KAO MEDIJANA (INTERKVARTILNI OPSEG). PRETHODNO ISKUSTVO JE OZNAČENO KAO: 'F' = ISPITANICI SA DUGOGODIŠNJIM ISKUSTVOM U ELEKTROTAKTILNOM *FEEDBACK*-U; 'S' = ISPITANICI SA ISKUSTVOM U NEKOM VIDU ELEKTRIČNE STIMULACIJE; 'B' = ISPITANICI BEZ IKAKVOG ISKUSTVA

Ispitanici	Test 1 – mekoća [%]	Test 2 – veličina [%]	Test 3 – oboje [%]	Test 3 – marginalna mekoća [%]	Test 3 – marginalna veličina [%]	Trajanje familiarizacije [min]	Trajanje učenja sa podsticanjem [min]	Prethodno iskustvo (F/S/B)
ISP1	100	90	100	100	100	1.4	2.9	S
ISP2	100	100	83.3	83.3	91.7	2.0	1.4	S
ISP3	50	100	75	75	100	2.4	1.5	S
ISP4	100	100	100	100	100	3.4	2.3	F
ISP5	100	100	100	100	100	1.5	1.9	F
ISP6	100	90	75	75	100	1.9	5.2	S
ISP7	100	100	100	100	100	1.1	1.2	F
ISP8	60	100	66.7	66.7	100	4.9	4.6	B
ISP9	100	100	100	100	100	6.3	1.1	B
ISP10	80	100	91.7	91.7	100	2.4	2.5	S
Mediana (IKO)	100 (20)	100 (0)	96 (25)	96 (25)	100 (0)	2.2 (1.9)	2.1 (1.5)	

Tvrdnju 1 ocjenom 7 i postigao SUP mekoće jednaku 60% i 66,7% u testu 1 i testu 3, respektivno. To ukazuje na to da su ispitanici bili svjesni u kojoj mjeri su se elektrotaktilne informacije dobro razumjele, a kada su bile nejasne i nisu bili sigurni u svoj odgovor.

Nemoguće je izvući jasne zaključke o direktnom odnosu između dužine trajanja treninga i pojedinačne SUP postignute u eksperimentu. Neki ispitanici koji su duže trenirali postigli su odlične rezultate, dok su se neki izjednačili ili bili gori od subjekata koji su dosta kraće trenirali. Međutim, primjetno je da su ispitanici bez prethodnog iskustva u bilo kojoj vrsti električne stimulacije (ISP8 i ISP9) imali najduže trajanje treninga (Tabela II). Štaviše, ispitanici sa dugogodišnjim iskustvom (ISP4, ISP5 i ISP7) uglavnom su se mogli pripremiti za test nakon kratkog treninga (do 5 minuta). Rezultati upitnika pokazuju da su i iskusni i neiskusni subjekti istakli važnost treninga (Tvrdnja 3, Tabela III).

Uzimajući u obzir da su sva tri ispitanika sa dugogodišnjim iskustvom imala ukupnu SUP jednaku 100%, možemo zaključiti da rezultati u velikoj mjeri zavise od prethodnog iskustvu ispitanika u elektrotaktilnom *feedback*-u, što je u skladu s našim prethodnim istraživanjem koje je pokazalo dugoročni efekat učenja [25]. Međutim, treba skrenuti pažnju na to da je ISP8 postigao identičan SUP u prepoznavanju veličine lopte bez ikakvog prethodnog iskustva, što ukazuje na to da su poruke bile intuitivne i da ih je lako bilo naučiti nakon kratkog treninga (manje od 10 minuta). S druge strane, bilo je i ispitanika koji, uprkos protokolu koji je obuhvatao prenos samo šest različitih poruka i činjenici su imali neograničeno vrijeme za trening, nisu mogli postići visoku SUP. To se može pripisati drugim faktorima poput taktike učenja ispitanika, brzine učenja, mentalne koncentracije, stope navikavanja, fizičkog stanja (poput umora) i sl. koji mogu da igraju važnu ulogu. U literaturi [12], [25] postoje naznake da se SUP stimulusa povećava tokom treninga, ali da bi se taj i slični fenomeni ispitali, trebalo bi provesti dalje eksperimente koji bi uključili veći broj ponavljanja, duži i trening s više pokušaja.

TABELA III

REZULTATI UPITNIKA O POSTIGNUTIM PERFORMANSAMA PRILIKOM PREPOZNAVANJA MEKOĆE I VELIČINE (TVRDNJA 1 I TVRDNJA 2) I VAŽNOSTI TRENINGA (TVRDNJA 3)

	Tvrdnja 1	Tvrdnja 2	Tvrdnja 3
ISP1	10	8	10
ISP2	8	10	10
ISP3	6	10	9
ISP4	10	9	8
ISP5	10	10	10
ISP6	8	10	9
ISP7	9	9	9
ISP8	7	10	9
ISP9	10	10	8
ISP10	7	9	7

VI. ZAKLJUČAK

Cilj ove studije bio je ispitati mogućnost pružanja povratnih informacija sa telemanipulisanog aktuatora promjenjive krutosti o veličini i mekoći uhvaćenog predmeta putem elektrotaktilnog sistema za stimulaciju na vrhu prsta. U tu svrhu izvedeni su eksperimenti na Qbmove Maker Pro aktuatoru, dok je elektrotaktilni *feedback* obezbjeđen zahvaljujući programibilnom stimulatoru i specijalno dizajniranoj, fleksibilnoj, površinskoj, *multi-pad* elektrodi.

Aktuator promjenjive krutosti omogućio je izvođenje eksperimenta s različitim osobinama predmeta. Uz odgovarajuće podešavanje krutosti, mogla bi se postići osjetljivost na meke ili tvrde materijale i u oba slučaja operatoru pružiti raspoznatljiva povratna informacija. Veličina i mekoća predmeta procjenjivani su na osnovu položaja hvataljke pri kontaktu sa objektom, i podataka o momentu sile izmjerenim senzorima unutar aktuatora, a kodirani su prostornom (tj. promjenom aktivnog polja elektrode) i frekvencijskom modulacijom.

Testovi su pokazali da se tri nivoa frekvencije stimulacije (koja predstavljaju mekoću objekta) i dvije lokacije stimulusa (koje predstavljaju veličinu objekta) mogu prepoznati sa prosječnim SUP jednakom 100 (20)% i 100 (0)%, respektivno. U kombinovanom frekvencijskom i prostornom kodiranju, prosječna SUP bila je 96 (25)%, i ne pokazuje statistički značajne razlike sa SUP osobina objekata u pojedinačnim testovima (test 1 i test 2). Preliminarni rezultati predstavljeni u ovom radu ohrabruju dalja istraživanja primjene elektrotaktilnog *feedback*-a u robotskim sistemima za telemanipulaciju i drugim scenarijima od interesa.

ZAHVALNICA

Rad u ovoj studiji izveden je u okviru projekta TACTILITY, koji je finansiran od strane Evropske Unije H2020-ICT-2018-2020/H2020-ICT-2018-3 prema sporazumu o dodjeli bespovratnih sredstava broj 856718. Ovom prilikom, autori bi se željeli zahvaliti svim partnerima u projektu na radu koji nam je omogućio predstavljeno istraživanje. Zahvaljujemo se svim zdravim volonterima koji su učestvovali u našoj studiji.

LITERATURA

[1] J. Van Oosterhout, D. A. Abbink, J. F. Koning, H. Boessenkool, J. G. W. Wildenbeest, and C. J. M. Heemskerk, “Haptic shared control improves hot cell remote handling despite controller inaccuracies,” *Fusion Eng. Des.*, vol. 88, no. 9–10, pp. 2119–2122, 2013.

[2] T. B. Sheridan, “Space teleoperation through time delay: Review and prognosis,” *IEEE Trans. Robot. Autom.*, vol. 9, no. 5, pp. 592–606, 1993.

[3] C.-H. King *et al.*, “Tactile feedback induces reduced grasping force in robot-assisted surgery,” *IEEE Trans. Haptics*, vol. 2, no. 2, pp. 103–110, 2009.

[4] M. Ouhyoung, W.-N. Tsai, M.-C. Tsai, J.-R. Wu, C.-H. Huang, and T.-J. Yang, “A low-cost force feedback joystick and its use in PC video games,” *IEEE Trans. Consum. Electron.*, vol. 41, no. 3, pp. 787–794, 1995.

[5] Y.-Y. Liao, L.-R. Chou, T.-J. Horng, Y.-Y. Luo, K.-Y. Young, and S.-F. Su, “Force reflection and manipulation for a VR-based telerobotic system,” *Proc.-Natl. Sci. Counc. Repub. China Part Phys. Sci. Eng.*, vol. 24, no. 5, pp. 382–389, 2000.

[6] M. Zhou, S. Tse, A. Derevianko, D. B. Jones, S. D. Schwaitzberg, and C. G. L. Cao, “Effect of haptic feedback in laparoscopic surgery skill acquisition,” *Surg. Endosc.*, vol. 26, no. 4, pp. 1128–1134, 2012.

[7] G. Tholey, J. P. Desai, and A. E. Castellanos, “Force feedback plays a significant role in minimally invasive surgery: results and analysis,” *Ann. Surg.*, vol. 241, no. 1, p. 102, 2005.

[8] C. R. Wagner, R. D. Howe, and N. Stylopoulos, “The role of force feedback in surgery: analysis of blunt dissection,” in *Haptic Interfaces for Virtual Environment and Teleoperator Systems, International Symposium on*, 2002, pp. 73–73.

[9] H. Kim, C. Seo, J. Lee, J. Ryu, S. Yu, and S. Lee, “Vibrotactile display for driving safety information,” in *2006 IEEE Intelligent Transportation Systems Conference*, 2006, pp. 573–577.

[10] I. Zubrycki and G. Granosik, “Novel haptic device using jamming principle for providing kinaesthetic feedback in glove-based control interface,” *J. Intell. Robot. Syst.*, vol. 85, no. 3–4, pp. 413–429, 2017.

[11] A. Y. Szeto and F. A. Saunders, “Electrocutaneous stimulation for sensory communication in rehabilitation engineering,” *IEEE Trans. Biomed. Eng.*, no. 4, pp. 300–308, 1982.

[12] M. Štrbac *et al.*, “Integrated and flexible multichannel interface for electrotactile stimulation,” *J. Neural Eng.*, vol. 13, no. 4, p. 046014, 2016.

[13] M. Isaković *et al.*, “Electrotactile feedback improves performance and facilitates learning in the routine grasping task,” *Eur. J. Transl. Myol.*, vol. 26, no. 3, 2016.

[14] S. Dosen *et al.*, “Multichannel electrotactile feedback with spatial and mixed coding for closed-loop control of grasping force in hand prostheses,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 3, pp. 183–195, 2016.

[15] G. K. Patel, S. Dosen, C. Castellini, and D. Farina, “Multichannel electrotactile feedback for simultaneous and proportional myoelectric control,” *J. Neural Eng.*, vol. 13, no. 5, p. 056015, Oct. 2016, doi: 10.1088/1741-2560/13/5/056015.

[16] D. S. Pamungkas and K. Ward, “Electro-Tactile Feedback System to Enhance Virtual Reality Experience,” *Int. J. Comput. Theory Eng.*, vol. 8, no. 6, pp. 465–470, Dec. 2016, doi: 10.7763/IJCTE.2016.V8.1090.

[17] D. S. Pamungkas and K. Ward, “Tele-operation of a robot arm with electro tactile feedback,” in *2013 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, Wollongong, NSW, Jul. 2013, pp. 704–709. doi: 10.1109/AIM.2013.6584175.

[18] D. S. Pamungkas and K. Ward, “Electro-tactile feedback system for achieving embodiment in a tele-operated robot,” in *2014 13th International Conference on Control Automation Robotics & Vision (ICARCV)*, Singapore, Dec. 2014, pp. 1706–1711. doi: 10.1109/ICARCV.2014.7064573.

[19] D. Pamungkas and K. Ward, “Immersive teleoperation of a robot arm using electro-tactile feedback,” in *2015 6th International Conference on Automation, Robotics and Applications (ICARA)*, Queenstown, New Zealand, Feb. 2015, pp. 300–305. doi: 10.1109/ICARA.2015.7081164.

[20] K. Li, P. Boyd, Y. Zhou, Z. Ju, and H. Liu, “Electrotactile Feedback in a Virtual Hand Rehabilitation Platform: Evaluation and Implementation,” *IEEE Trans. Autom. Sci. Eng.*, vol. 16, no. 4, pp. 1556–1565, Oct. 2019, doi: 10.1109/TASE.2018.2882465.

[21] T. J. Arakeri, B. A. Hasse, and A. J. Fuglevand, “Object discrimination using electrotactile feedback,” *J. Neural Eng.*, vol. 15, no. 4, p. 046007, Aug. 2018, doi: 10.1088/1741-2552/aabc9a.

[22] G. Chai, J. Briand, S. Su, X. Sheng, and X. Zhu, “Electrotactile Feedback with Spatial and Mixed Coding for Object Identification and Closed-loop Control of Grasping Force in Myoelectric Prostheses,” in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Berlin, Germany, Jul. 2019, pp. 1805–1808. doi: 10.1109/EMBC.2019.8856508.

[23] “qbrobotics: robotics and automation.” https://qbrobotics.com/ (accessed Feb. 02, 2021).

[24] “45c5a1_792590e00b134129b2b6363a1ea7de45.pdf.” Accessed: Feb. 02, 2021. [Online]. Available: https://www.qbrobotics.com/wp-content/uploads/2016/03/45c5a1_792590e00b134129b2b6363a1ea7de45.pdf

[25] M. Strbac *et al.*, “Short- and Long-Term Learning of Feedforward Control of a Myoelectric Prosthesis with Sensory Feedback by Amputees,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 11, pp. 2133–2145, Nov. 2017, doi: 10.1109/TNSRE.2017.2712287.

ABSTRACT

The major precondition to take full advantage of teleoperation systems is a bidirectional exchange of haptic information between an operator and the remote environment, enabling the operator to perceive collisions, contact forces, weight, object shapes, surface textures, etc. To address some of these requirements, we propose the use of electrotactile stimulation with spatial and frequency coding of information. The envisioned system is comprised of a remote compliant robot actuator, stimulator unit and a surface multi-pad electrode that is placed on the subject’s index fingertip. Unlike similar haptic feedback setups, in our study the subjects relied exclusively on haptic feedback, without visual or auditory cues. Experimental results showed that tactile stimulation can serve as feedback regarding the softness (90% average recognition rate of 3 softness levels) and the size (98% average recognition rate of 2 sizes) of the object grasped by a remote actuator.

Electrotactile feedback on object properties manipulated by a compliant robotic acturator

Gorana Marković, Jovana Malešević, Milica Isaković, Miloš Kostić, Matija Štrbac, Kosta Jovanović

Pronalazak Optimizacione Funkcije Kretanja iz Simulirane Demonstracije Pokreta Čučnja

Filip Bečanović, Član, IEEE, Vincent Bonnet, Član, IEEE,
Samer Mohammed, Član, IEEE, Kosta Jovanović, Član, IEEE

Sažetak—U ovom radu, posebno je uspostavljen pojednostavljeni ravanski biomehanički model čoveka za zadatak čučnja. Koristeći taj model, generišu se višestruki optimalni pokreti u odnosu na različite biomehaničke optimizacione funkcije i njihove linearne kombinacije. Iz optimalnih pokreta, optimizacione funkcije kojim su generisani pronalaze se pomoću inverzne optimizacije. Optimizaciona funkcija koja daje kvalitativno najsličnije kretanje ljudskom je kombinacija minimalnog zglobnog momenta, ubrzanja i snage. Medjutim, čini se da stvarno ljudsko kretanje minimizira i neke optimizacione funkcije koje nisu razmatrane u ovom radu.

I. Uvod

Ideja generisanja ljudskih pokreta optimizacijom je zastupljena u literaturi. Pregled takvih pristupa za pokret ustajanja sa stolice i dizanja kutije dat je u [1]. Unutar [2], autor koristi planarni model sa 5 stepeni slobode kako bi generisao zglobne trajektorije za dizanje kutije sa minimalnim zglobnim momentom u članku, i pokazuje da to dovodi do stabilnosti modela po principu Tačke Nultog Momenta. U [3], autori koriste trodimenzionalni model sa 55 stepeni slobode da bi generisali optimalne trajektorije za pokret dizanja kutije, koristeći dve optimizacione funkcije.

Ideja modelovanja ljudskog kretanja kao optimizacije, zatim pronalaska optimizacione funkcije isto je obradjena u literaturi. Unutar [4], autori rada modeluju ljudsko kretanje unutar prostorije kao kretanje materijalne tačke u dvodimenzinalnom prostoru (iz ptičije perspektive), a izbor trajektorije modeluju kao optimizacioni proces. Uz pomoć takozvane "bi-level" metode za inverznu optimizaciju, oni uspevaju da pronadju optimizacionu funkciju koja generiše tajektorije koje su najpodudarnije sa snimljenim podacima pravih ljudskih trajektorija. U [5] koristeći se sličnom metodom, autori uporedjuju rezultate simulacije sa ljudskim podacima za pokret hvata rukom u transverzalnoj ravni. Dok u [6], izučavan pokret je hvat rukom u sagitalnoj ravni.

U [7], predstavljena je nova metoda za inverznu optimizaciju zasnovana na zadovoljavanju Karuš-Kun-Takerovih (KKT) uslova. Ova metoda iskorišćena je u [8] za izučavanje pokreta otvaranja vrata, ili izvlačenja i zatvaranja fioke.

U ovom radu izučavaćemo pokret čučnja, kao jednostavan simetričan pokret koji se može posmatrati samo u sagitalnoj ravni ljudskog tela, i koji je odredjen većinski zglobovima donjeg dela tela. Upotrebićemo splajn reprezentaciju za trajektorije kao što se to radi u robotici [9], upotrebićemo standardnu formulaciju optimizacionog problema za generisanje trajektorija, a za inverznu optimizaciju upotrebićemo KKT metodu.

II. Biomehanički model

Ova deonica predstaviće modelovanje ljudskog tela za pokret dizanja kutije.

Nalik robotici, ljudsko telo je modelovano kao uzastopna veza krutih tela spojenih zglobovima. Da bi se uprostio model, iskrorišćena je simetrija zadatka kako bi se apstrahovala kretnja i svela na ravansku kretnju u sagitalnoj ravni ljudskog tela. Tim postupkom dobijamo model nalik onome na Slici 1. Osnova ljudksog modela je stopalo, dok segmente čine potkolenica, butina, i skup svih delova tela iznad trupa. Istim redosledom, povezuju ih članak, koleno, i kuk. Zajedno, ovi zglobovi daruju modelu tri ravanska stepena slobode. Treba primetiti da zbog ravanske abstrakcije, simetrični delovi tela bivaju duplirani, što znači da se noge i njihovi zglobovi svode na jednu nogu sa istim zglobovima.

Visina i masa modela fiksirane su, i jednake 1.83m i 81kg respektivno. Kinematski i inercijalni parametri delova tela odredjeni su korišćenjem antropometričkih tabela iz [10], iz kojih se mogu odrediti dužine segmenata, masa segmenata, pozicije centara mase segmenata, kao i matrice inercije segmenata. Duplirani segmenti modela imaju duplirane mase kao i duplirane inercije. Istovremeno, sve iznad trupa je spojeno u jedan segment, a samim tim spojene su njihove mase i inercije.

S obzirom na to da ćemo izučavati kinematiku zglobova u zglobnom prostoru, kao i kinematiku završnog uredjaja (ruke) u Kartezijanskom prostoru, ali i dinamiku u zglobnom prostoru (momenti sile u zglobovima), kao

Filip Bečanović is a joint Ph.D. student with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia, and the University of Paris-Est Créteil, Laboratory of Images, Signals and Intelligent Systems, 120 rue Paul Armangot, 94400 Vitry sur Seine, (email: becaphilippe@gmail.com).

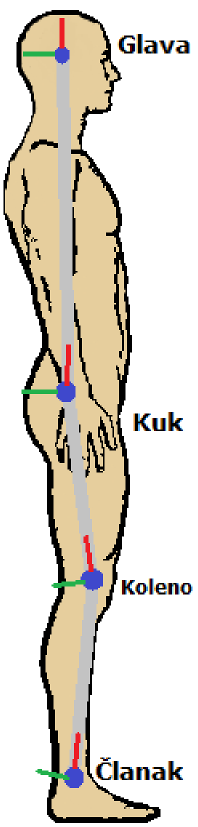
Vincent Bonnet is with the Laboratory for Analysis and Architecture of Systems, 7 avenue du Colonel Roche, 31031 Toulouse, France, (email: bonnet.vincent@gmail.com).

Samer Mohammed is with the University of Paris-Est Créteil, Laboratory of Images, Signals and Intelligent Systems, 120 rue Paul Armangot, 94400 Vitry sur Seine, (email: samer.mohammed@u-pec.fr).

Kosta Jovanović is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia, (email: kostaj@etf.rs).

i eksterne sile i momente u Kartezijanskom prostoru, potrebno je postaviti direktan kinematski, kao i inverzni dinamički model. Direktan kinematski model, koji mapira ugaone položaje zglobova u Kartezijanski položaj završnog uredjaja, može se lako proračunati korišćenjem takozvanih proksimalnih Denavit-Hartenberg parametara [11]. Za inverznu dinamiku koja mapira zglobne ugaone pozicije, brzine, ubrzanja, zajedno sa eksternim silama i momentima, u zglobne momente, korišćen je softverski paket SYMORO+ [12] za automatsko generisanje koda.

Ograničavajući zglobni uglovi kao i maksimalni zglobni momenti koji su neophodni za kasnije proračune, određeni su iz [13].



Slika 1. Ravanski biomehanički model čoveka sa tri stepena slobode

III. Kvintni Splajn

Ova deonica diskutovaće splajn predstavu [14] zglobnih trajektorija.

Pokret čučnja može se zamisliti kao da model počinje iz stajaćeg položaja, zatim spušta kuk ispod odredjene visine, i vraća se u stajaći položaj. Naša namera je da napravimo algoritam za planiranje pokreta od početnog do krajnjeg položaja, tako da poštuje odredjena ograničenja i ispunjava odredjene zahteve, istovremeno minimizirajući neku složenu optimizacionu funkciju.

Kako bismo to uspeli, neophodna nam je matematička predstava za zglobne putanje svih zglobova. Ona mora biti istovremeno dovoljno ekspresivna kako bi mogla predstaviti sve detalje jedne složene kretnje, ali ne isuviše kompleksna kako bi vreme proračunavanja bili relativno kratko. U ovom radu, opredelili smo se za kvintnu splajn predstavu.

Pretpostavljajući da je q funkcija vremena koja predstavlja trajektoriju jednog zgloba,

$$q \in \{q_{Clanak}, q_{Koleno}, q_{Kuk}\}$$

može se pretstaviti preko N parova vreme-zglobni ugao, koje nazivamo putne tačke:

$$(t_0, q_0^*), (t_1, q_1^*), \ldots, (t_N, q_N^*). \tag{1}$$

Ukoliko je u početnom i krajnjem trenutku čovek u stanju mirovanja[1] postoji jedinstvena deo-po-deo polinomska trajektorija petog stepena (kvintna), koja prolazi kroz sve putne tačke i zadovoljava odredjene uslove glatkosti. Tajektorija se modeluje, između vremena t_i i t_{i+1}, kao polinom 5-og stepena izražen na sledeći način

$$q_i(t) = a_i + b_i t + c_i t^2 + d_i t^3 + e_i t^4 + f_i t^5. \tag{2}$$

Celokupna trajektorija je deo-po-deo kvintna funkcija koja je sastavljena od sleda pojedinačnih kvintnih polinomskih trajektorija.

$$q(t) = \begin{cases} q_0(t), & \text{if } t_0 \le t < t_1 \\ \vdots & \\ q_{N-1}(t), & \text{if } t_{N-1} \le t < t_N \end{cases} \tag{3}$$

Ukoliko je dat skup putnih tačaka, postoji jedinstveni skup koeficijenata $\{a_i, b_i, c_i, d_i, e_i, f_i\}_{i=0,\ldots,N-1}$ takvih da konačna trajektorija prolazi kroz sve putne tačke i jeste neprekidna, zajedno sa svojih prvih četiri izvoda, u intervalu vremena (t_0, t_N).

Da bi se jedinstveno izračunali ovih $6N$ koeficijenata, možemo postaviti isti broj linearnih jednačina koje ih sadrže. Da bismo osigurali da svi kvintni polinomi prolaze kroz putne tačke na njihovim krajevima, izjednačavamo ih sa vrednošću funkcije u tim putnim tačkama.

$$q_i(t_i) = q_i^* \tag{4}$$
$$q_i(t_{i+1}) = q_{i+1}^* \tag{5}$$

Svaki pojedinačan kvintni polinom mora proći kroz putne tačke na sopstvenim krajevima, što znači da jednačine (4) i (5) važe za $i = 0, \ldots, N-1$, i samim tim su odgovorne za $2N$ jednačina.

Neprekidnost izvoda trajektorije u unutrašnjim putnim tačkama može se osigurati izjednačavanjem izvoda susednih kvintnih polinoma u njihovoj zajedničkoj putnoj tački.

$$\dot{q}_i(t_{i+1}) = \dot{q}_{i+1}(t_{i+1}) \tag{6}$$
$$\ddot{q}_i(t_{i+1}) = \ddot{q}_{i+1}(t_{i+1}) \tag{7}$$
$$\dddot{q}_i(t_{i+1}) = \dddot{q}_{i+1}(t_{i+1}) \tag{8}$$
$$q_i^{(4)}(t_{i+1}) = q_{i+1}^{(4)}(t_{i+1}) \tag{9}$$

[1]Početna i krajnja brzina, kao i ubrzanje, su jednaki nuli.

S obzirom da (6), (7), (8), i (9) važe za $i = 1, \ldots, N-1$, i zbog toga su odgovorne za $4(N-1)$ jednačina.

Ovom sistemu linearnih jednačina nedostaje samo 4 jednačine da bi bio kvadratan. Preostale jednačine će podstaći to da model čoveka bude u stanju mirovanja na početku i kraju pokreta.

$$\dot{q}_0(t_0) = 0 \tag{10}$$

$$\ddot{q}_0(t_0) = 0 \tag{11}$$

$$\dot{q}_{N-1}(t_N) = 0 \tag{12}$$

$$\ddot{q}_{N-1}(t_N) = 0 \tag{13}$$

Rešavanjem ovog sistema od $6N$ jednačina sa $6N$ nepoznatih dobija se jedinsven skup koeficijenata, dakle i jedinstvena deo-po-deo kvintna trajektorija. Štaviše, ovi koeficijenti će linearno zavisiti od vrednosti funkcije u putnim tačkama.

Takodje je vredno pomenuti da izračunavanjem funkcije ili neke od njenih izvoda u nekom vremenu t' takvom da je $t_i < t' < t_{i+1}$, dobija se vrednost koja je linearno zavisna od koeficijenata i-tog polinoma, a samim tim i od vrednosti funkcije u putnim tačkama.

$$q(t') = a_i + b_i t' + c_i t'^2 + d_i t'^3 + e_i t'^4 + f_i t'^5 \tag{14}$$

Može se zaključiti da će vrednost funkcije $q(t')$ u bilo kom trenutku u vremenu $t_0 < t' < t_N$ biti linearno zavisna od vrednosti funkcije u putnim tačkama.

IV. Optimizacija

U ovoj deonici biće opisan optimizacioni problem koji se rešava prilikom generisanja trajektorija.

Kao što je pokazano u Deonici III, trajektorija svakog zgloba se može predstaviti korišćenjem N putnih tačaka. Pretpostavljajući sada da imamo trodimenzionalnu trajektoriju

$$q = \left[q_{Clanak}, q_{Koleno}, q_{Kuk}\right]^T$$

i da su putne tačke date za takvu trajektoriju, slično tome kako su date u (1), možemo ih upakovati u optimizacioni vektor

$$x = \left[q_0^{*T} \quad q_1^{*T} \quad \ldots \quad q_N^{*T}\right]^T$$

gde je svaka od njih vektor $q_i^* \in \mathbb{R}^3,\ i = 0, \ldots, N$.

Problem pronalaženja najprikadnijih zglobnih trajektorija u skladu sa nekom optimizacionom funkcijom, podešavajući pozicije putnih tačaka u prostoru zglobova, može se formalno predstaviti kao

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & h(x) = 0 \\ & g(x) \leq 0 \end{aligned} \tag{15}$$

gde je $x \in \mathbb{R}^{3(N+1)}$ optimizaciona promenljiva koja predstavlja vektor putnih tačaka, $f(x) : \mathbb{R}^{3(N+1)} \mapsto \mathbb{R}$ je optimizaciona funkcija, $h(x) : \mathbb{R}^{3(N+1)} \mapsto \mathbb{R}^m$ predstavlja orgraničenja u vidu sistema jednakosti, a $g(x) : \mathbb{R}^{3(N+1)} \mapsto \mathbb{R}^p$ predstavlja ograničenja u vidu sistema nejednakosti.

U ovom radu, optimizaciona funkcija predstavlja linearnu kombinaciju skupa optimizacionih funkcija definisane vektorskom funkcijom $\tilde{f}(x)$. Optimizaciona funkcija može se izraziti kao

$$f(x) = \sum_{i=1}^{N_{CF}} \alpha_i f_i(x) = \alpha^T \tilde{f}(x). \tag{16}$$

Gde koeficijenti moraju biti pozitivni $\alpha_i \geq 0,\ i = 1, \ldots, N_{CF}$ kako bi se održala lokalna konveksnost funkcija, kao i normalizovani $\sum_{i=1}^{N_{CF}} \alpha_i = 1$ kako bi svakom rešenju odgovarala jedinstvena linearna kombinacija.

Skup otpimizacionih funkcija koje se razmatraju u ovom radu date su u Tabeli I. Dok su ograničenja predstavljena u Tabeli II.

Vrednosti trajektorije, brzine i ubrzanja u pojedinim trenucima $(q(t'), \dot{q}(t'), \ddot{q}(t'))$ mogu se dobiti iz koeficijenata dobijenih pomoću procedure opisane u Deonici III. Kartezijanska pozicija kuka (x_{Kuk}, y_{Kuk}) može se dobiti pomoću direktnog kinematskog modela. Zglobni momenti Γ, mogu se dobiti pomoću inverznog dinamičkog modela, kao i centar pritiska COP [15].

Tabela I
Optimizacione funkcije

Ime	Izraz
Zbir Kvadrata Zglobnih Momenata	$f_1(x) = \sum_{j=1}^{6} \sum_{i=0}^{N_S} \Gamma_{ji}^2$
Zbir Kvadrata Zglobnih Ubrzanja	$f_2(x) = \sum_{j=1}^{6} \sum_{i=0}^{N_S} \ddot{q}_{ji}^2$
Zbir Kvadrata Zglobnih Snaga	$f_4(x) = \sum_{j=1}^{6} \sum_{i=0}^{N_S} (\Gamma_{ji} \cdot \dot{q}_{ji})^2$

Tabela II
Funkcije ograničenja

Ime	Izraz
Sistemi nejednakosti	
Uslov Stabilnosti	$x_{peta} \leq COP(t_i) \leq x_{prst},\ i = 0, ..., N_{itp}$
Ograničenja Zglobnog Momenta	$\Gamma_{min} \leq \Gamma(t_i) \leq \Gamma_{max},\ i = 0, ..., N_{itp}$
Uslov Čučnja	$y_{Kuk}(t_i) \leq y_{Prag},\ i = 0, ..., N_{itp}$
Zglobna Ograničenja	$q_{min} \leq q(t_i) \leq q_{max},\ i = 0, ..., N_{itp}$
Sistemi jednakosti	
Početni Uslovi	$q(t_0) = q_{uspravno}$
Krajnji Uslovi	$q(t_N) = q_{uspravno}$

V. Inverzna Optimizacija

U ovoj deonici biće opisano kako izvršiti inverznu optimizaciju na osnovu Karuš-Kun-Takerovih uslova optimalnosti [7].

Ako je dat rezultat x^* optimizacionog problema (15) gde je optimizaciona funkcija linearna kombinacija različitih funkcija kao u (16), cilj inverzne optimizacije jeste da se pronadju koeficijenti α_i, $i = 1, \ldots, N_{CF}$ te linearne kombinacije. Pretpostavljajući da je data snimljena ljudska trajektorija u obliku putnih tački, inverzna optimizacija bi trebalo da odredi koje optimizacione funkcije je čovek koristio prilikom generisanja takve trajektorije, i time nam omogućiti generisanje čovekolikih trajektorija u budućnosti, korišćenjem pronađene optimizacione funkcije.

Koristićemo KKT uslove da rešimo problem inverzne optimizacije. Za svaki optimizaconi problem izražen kao u (15), možemo definisati Lagranžijan uz pomoć promenljivih λ i μ koje nazivamo Lagranžovim multiplikatorima.

$$L(x, \lambda, \mu) = f(x) + \lambda^T \cdot h(x) + \mu^T \cdot g(x) \tag{17}$$

Poznato je da optimalno rešenje x^* takozvanog primalnog problema (15) mora imati komplementarno dualno rešenje (λ^*, μ^*). Zajedno, oni moraju zadovoljavati KKT uslove koji su dati u jednačinama (18)-(22).

$$\begin{aligned} \nabla_x L(x^*, \lambda^*, \mu^*) &= 0 && (18)\\ h(x^*) &= 0 && (19)\\ g(x^*) &\leq 0 && (20)\\ \mu_i^* g_i(x^*) &= 0 \qquad i = 1, \ldots, p && (21)\\ \mu^* &\geq 0 && (22) \end{aligned}$$

S obzirom na to da je naša optimizaciona funkcija linearna kombinacija skupa funkcija, možemo Lagranžijan izraziti na sledeći način, i uočiti njegovu zavisnost od koeficijenata α koje želimo povratiti.

$$L(\alpha, x, \lambda, \mu) = \alpha^T \cdot \tilde{f}(x) + \lambda^T \cdot h(x) + \mu^T \cdot g(x) \tag{23}$$

Dakle KKT uslovi zavise od koeficijenata α ali takodje i od dualnih promenljivih (λ^*, μ^*) koje nisu date. Naš pristup povratku koeficijenata α će se zasnivati na zadovoljavanju KKT uslova, a zahtevaće i da pored koeficijenata tražimo vrednost multiplikatora (λ^*, μ^*).

Definisaćemo reziduale kao što je uradjeno u [16].

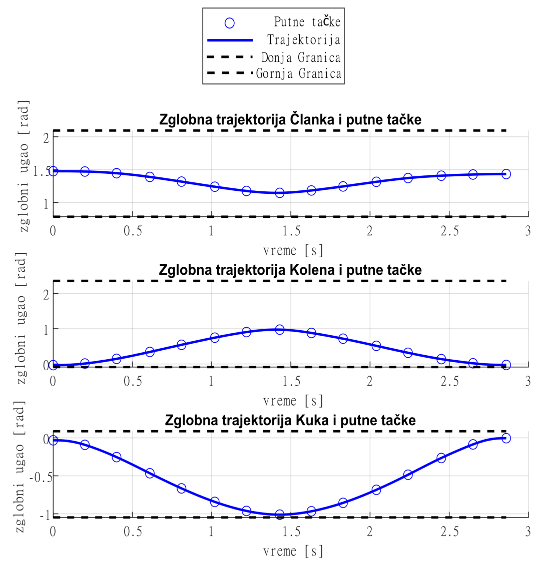
$$\begin{aligned} r_{stat} &= \nabla_x L(\alpha, x^*, \lambda^*, \mu^*) && (24)\\ r_{comp} &= \mathrm{diag}(g(x^*))\mu^* && (25) \end{aligned}$$

Zatim možemo formulisati želju za zadovoljavanjem KKT uslova kao optimizacioni problem pretrage za promenljivima $(\alpha, \lambda^*, \mu^*)$ takvim da su reziduali iz (24) i (25), koji odgovaraju uslovima (18) i (21), što bliži nuli a da istovremeno poštuju uslov (22) kao i uslov pozitivnosti i normalizacije koeficijenata α.

$$\begin{aligned} \min_{(\alpha,\lambda,\mu)} \quad & \|r_{stat}\|_2^2 + \|r_{comp}\|_2^2 \\ \text{s.t.} \quad & \alpha \geq 0 \\ & \mu \geq 0 \\ & \sum_{i=1}^{N_{CF}} \alpha_i = 1 \end{aligned} \tag{26}$$

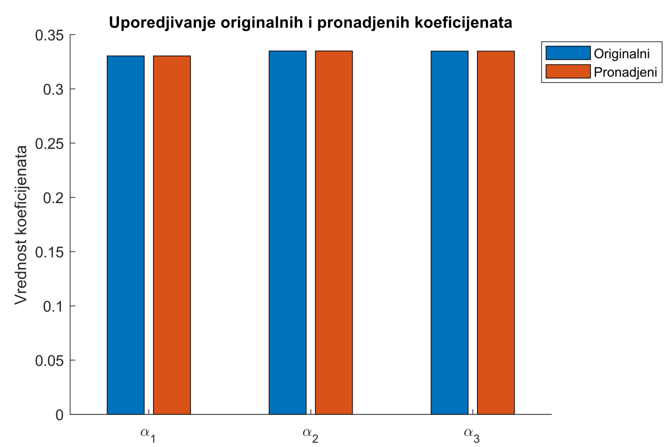
VI. Rezultati

Optimalne trajektorije generisane su po postupku opisanom u Deonici IV, sa $N = 15$ putnih tačaka, i sa koeficijentima $\alpha_1 = \frac{1}{3}$, $\alpha_2 = \frac{1}{3}$ i $\alpha_3 = \frac{1}{3}$, koji redom odgovaraju funkcijama zbira kvadrata zglobnih momenata, zbira kvadrata zglobnih ubrzanja, i zbira kvadrata globnih snaga. Dobijene optimalne putne tačke, kao i odgovarajuće trajektorije dobijene interpolacijom opisanom u Deonici III, prikazane su na Slici 2.



Slika 2. Rezultati optimizacije

Rezultati inverzne optimizacije nad trajektorijama, odnosno putnim tačkama, generisanim sa prethodno pomenutim koeficijentima prikazani su na Slici 3. Na slici su uporedjeni koeficijenti koji su iskorišćeni za proces direktne optimizacije sa pronadjenim koeficijenti metodom inverzne optimizacije opisane u Deonici V.



Slika 3. Rezultati inverzne optimizacije

Ova dva prikazana grafika predstavljaju uzorak preko koga želimo da pokažemo da metoda radi u simulaciji. Metoda je isprobana za 3D rešetku vrednosi koeficijenata $(\alpha_1, \alpha_2, \alpha_3)$ i pokazala je zadovoljavajuće rezultate na svim ispitanim vrednostima.

Generisanje optimalnih trajektorija u industrijskim zadacima potrebno je za procenu ergonomije tog pokreta kod ljudi, i samim tim odredjivanja nekih pravila izvodjenja pokreta kako bi se sprečile ili umanjile mišićno-skeletalne povrede.

Takodje je od značaja mogućnost odredjivanja principa po kojima se kreću zdravi ljudi, i uporedjivanja sa principima po kojima se kreću ljudi sa odredjenim patologijama, kako bi se mogle davati bolje smernice pacijentima prilikom rehabilitacije.

Budući radovi mogu primenjivati ovu metodu na pravim ljudskim podacima kako bi se pronašle odrednice kojima se ljudi vode prilikom generisanja sopstvenih pokreta. Ova metoda se takodje može primeniti i na drugim pokretima, medjutim morao bi se postaviti drugačiji optimizacioni model, sa ograničenjima koja bi zavisila od samog pokreta.

Zahvalnica

Ovo istraživanje podržano je od strane Ministarstva Inostranih Poslova Republike Francuske, i Ministarstva Prosvete, Nauke i Tehnološkog Razvoja Republike Srbije, unutar programa za bilateralnu naučnu saradnju izmedju Republike Francuske i Republike Srbije, projekat #17, HUMAN-COMAN. Ovo istraživanje takodje je podržano od strane Campus France, Francuske agencije za promociju visokog obrazovanja, i internacionalne mobilnosti.

Reference

[1] Erik A Chumacero-Polanco and James Yang. „A review on human motion prediction in sit to stand and lifting tasks". in: International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. volume 50077. American Society of Mechanical Engineers. 2016, V01AT02A066.

[2] Ali Leylavi Shoushtari. „What strategy central nervous system uses to perform a movement balanced? Biomechatronical simulation of human lifting". in: Appl Bionics Biomech 10 (2013), pages 113–124.

[3] Yujiang Xiang andothers. „Human lifting simulation using a multi-objective optimization approach". in: Multibody System Dynamics 23.4 (2010), pages 431–451.

[4] Katja Mombaur, Anh Truong and Jean-Paul Laumond. „From human to humanoid locomotion—an inverse optimal control approach". in: Autonomous robots 28.3 (2010), pages 369–383.

[5] Sebastian Albrecht, Marion Leibold and Michael Ulbrich. „A bilevel optimization approach to obtain optimal cost functions for human arm movements". in: Numerical Algebra, Control & Optimization 2.1 (2012), page 105.

[6] Bastien Berret andothers. „Evidence for composite cost functions in arm movement planning: an inverse optimal control approach". in: PLoS Comput Biol 7.10 (2011), e1002183.

[7] Arezou Keshavarz, Yang Wang and Stephen Boyd. „Imputing a convex objective function". in: 2011 IEEE international symposium on intelligent control. IEEE. 2011, pages 613–619.

[8] Peter Englert, Ngo Anh Vien and Marc Toussaint. „Inverse KKT: Learning cost functions of manipulation tasks from demonstrations". in: The International Journal of Robotics Research 36.13-14 (2017), pages 1474–1488.

[9] J Zico Kolter and Andrew Y Ng. „Task-space trajectories via cubic spline optimization". in: 2009 IEEE International Conference on Robotics and Automation. IEEE. 2009, pages 1675–1682.

[10] Raphaël Dumas and Janis Wojtusch. Estimation of the Body Segment Inertial Parameters for the Rigid Body Biomechanical Models Used in Motion Analysis. In: Müller B., Wolf S.(eds) Handbook of Human Motion. 2018.

[11] Harvey Lipkin. „A note on Denavit-Hartenberg notation in robotics". in: International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. volume 47446. 2005, pages 921–926.

[12] Wisama Khalil and Denis Creusot. „SYMORO+: a system for the symbolic modelling of robots". in: Robotica 15 (1997), pages 153–161.

[13] Thomas Robert, Julien Causse and Gérard Monnier. „Estimation of external contact loads using an inverse dynamics and optimization approach: general method and application to sit-to-stand maneuvers". in: Journal of biomechanics 46.13 (2013), pages 2220–2227.

[14] Sky McKinley and Megan Levine. „Cubic spline interpolation". in: College of the Redwoods 45.1 (1998), pages 1049–1060.

[15] Philippe Sardain and Guy Bessonnet. „Forces acting on a biped robot. Center of pressure-zero moment point". in: IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans 34.5 (2004), pages 630–637.

[16] Adina Panchea. „Inverse optimal control for redundant systems of biological motion". phdthesis. Université d'Orléans, 2015.

Workspace Analysis of a Collaborative Bi-manual Industrial Robotic System

Jovan Šumarac, Kosta Jovanović and Aleksandar Rodić

***Abstract*—Bi-manual manipulation has been a focus of extensive academic research and has found its uses in industry as well. The workspace assessment of a robot is one of the key parameters in robot consideration for commercial purposes. As well, it is essential for the research of a bi-manual robotic system that often tends to replace humans in bi-manual tasks or directly share the workspace with humans. The goal of this paper is to present a detailed workspace analysis of a dual-arm collaborative robot. The dual-arm collaborative robot has been developed at the Robotics Laboratory at "Mihajlo Pupin" Institute and it is briefly presented in the paper. The workspaces of particular robot arms on the dual-arm system, a shared workspace for bimanual operation, and a manipulability analysis are presented. The simulations have been performed in Matlab, whereas CoppeliaSim robot simulator has been used for the visualization of the results. The presented results are an essential point in consideration of optimal trajectory planning and bi-manual collaborative robot control.**

***Index Terms*—Robot workspace; Bi-manual manipulation; Collaborative robots**

I. INTRODUCTION

Bi-manual or dual-arm manipulation is increasingly applied in modern robotics, and it is gaining further momentum with the advances in collaborative robotics. Bi-manual robots' similarity to human body form together with the need to replace human auxiliary work, as well as their increased workspace and task range are some of the main reasons. Similarity to human form also means easier integration of bi-manual robots in different environments originally intended for human workers, as well as for tasks which require human robot physical interaction. The analysis of the complete expanded workspace of such a robot, both in terms of reachability and manipulability, is very important for defining and coordinating its tasks.

Since its humble beginnings in 1940s and 50s mostly for tele-operation tasks, the development and research of bi-manual robots slowed down and gave way to single arm robots. However it again gained popularity and advanced significantly since the 1990s [1]. The focus was not only on scientific research, but also on producing commercial bi-manual robotic system.

Over the years many bi-manual robots were developed either for research or commercially. A detailed overview of the various scientific bi-manual platforms is given in [1]. Of those, the most interesting is Rollin' Justin [2]. It is a mobile robotic system and research platform that allows implementation of sophisticated control algorithms and dexterous manipulation. It is a powerful upper body humanoid robot with torso and two lightweight robot arms with four finger hands. Its workspace is quite large; its arm span is 3000 mm and the torso can move about 600 mm to the front and 300 mm to the back. However, a downside to this is that the robot is required to be mounted on a very stable base with a large footprint to prevent it from falling over.

In recent years, using collaborative bi-manual robots commercially has been a growing trend. In [3] a good overview of commercial bi-manual robots is given. The first such robot was called Baxter, presented in 2012 by Rethink Robotics. Its manipulator consists of a head, a torso and two arms with 7 degrees of freedom each. Its arm span is about 2600 mm and there is significant overlap between workspaces of both arms. Most of this common workspace is directly in front of the robot as it has a rotational limit at the shoulder [4]. ABB has also developed a commercial dual-arm collaborative robot called IRB 14000 Yumi. It can collaborate with a human and is intended for assembly of small parts. Its arms have relatively smaller ranges, about 560 mm each, resulting in a 1200 mm span and a smaller workspace[5]. Another smaller commercial bi-manual robot was also developed and presented by Epson in 2018, called the WorkSense W-01 with a similar arm range and workspace as Yumi [6].

The bi-manual robot analyzed in this paper was developed and made at the Robotics Laboratory at "Mihajlo Pupin" Institute. This paper presents the robot and an analysis of its workspace. The workspace is analyzed for each of the arms, and shared workspace is presented as well. Finally a manipulability analysis is performed. The simulation was done in Matlab and the results presented in CoppeliaSim robot simulator.

II. WORKSPACE ANALYSIS OF A BI-MANUAL ROBOT

The robot presented and analyzed in this paper was custom developed at the Robotics Laboratory [7]. The idea was to design and develop a cloud-enabled industrial human-size service robot. One of the main goals was to make it very

Jovan Šumarac is with the Robotics Laboratory,"Mihajlo Pupin" Institute, University of Belgrade, Volgina 15, 11060 Belgrade, Serbia (e- mail: jovan.sumarac@pupin.rs)

Kosta Jovanović is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: kostaj@etf.bg.ac.rs)

Aleksandar Rodić is with the Robotics Laboratory, "Mihajlo Pupin" Institute, University of Belgrade, Volgina 15, 11060 Belgrade, Serbia (e-mail: aleksandar.rodic@pupin.rs)

versatile and applicable in different tasks. The goal was also to make it the first-generation intelligent industrial service robot, meaning it possesses a significant level of artificial intelligence (such as perception, task planning, decision making, etc.).

The robot consists of several functional modules. Its base is a motorized cart that serves as a mobile platform and allows the transportation of the bi-manual robotic system. The mechanical robot torso has 4 degrees of freedom and the robot arms are mounted on it. The arms are two UR5 lightweight industrial manipulators with 6 degrees of freedom by Universal Robots.

Each robot arm has several modules at its end effector including force/torque sensors, wrist cameras and industrial grippers. Each arm is mounted on the torso at an angle of about 15 degrees which results in a larger shared workspace. Fig. 1 shows the robot in virtual reality in CoppeliaSim robot manipulator.

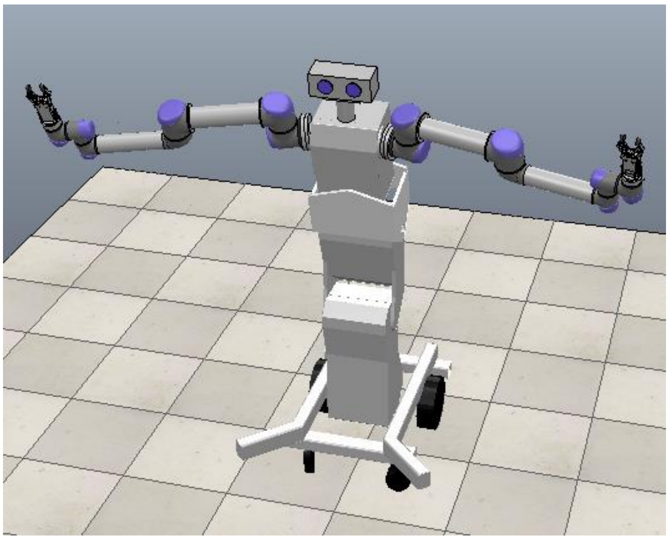


Fig. 1. The robotic system analyzed in this paper shown in a robot simulator

Fig. 1. shows the robot in a slightly bent forward torso configuration, which is common if the objects of interest are in front of the robot. It also shows the robot with a stylized head and eyes. This was done because the development of a vision and audio system is one of the goals for this robot as well.

The prototype of the mechanical torso currently developed at the Robotics Laboratory is shown in Fig. 2. The mechanical torso is of variable geometry. It can be in a standing position like in Fig. 2 which is its fully extended position. It can also be fully contracted e.g. in a bent position. In this paper the workspace analysis is performed for one, common configuration of the torso shown in Fig. 1. The results for the whole robot workspace for this torso configuration can easily be generalized for other torso positions.

A key issue in developing such a robotic system is defining and analyzing its workspace. Besides the obvious benefit of an increased workspace due to two robot arms, it is important to calculate the limits of the workspace in terms of reachability as well as identify problem areas within the workspace in terms of manipulability. This analysis is later used for designing the robots environment as well as the kind of tasks it will perform.

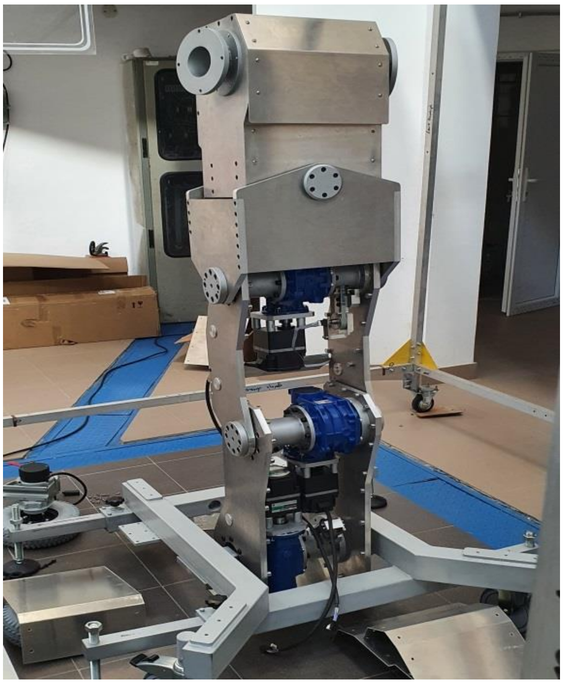


Fig. 2. The prototype of the mechanical torso of the analyzed robot.

Workspace analysis starts with the analysis of the individual workspaces of the robot arms. The arm ranges are provided by the manufacturer [8]. The idea behind the algorithm is to generate a relatively large number of random configurations within joint limits for both of the arms. Forward kinematics is then calculated to get the Cartesian coordinates of the end effector, and save its position for later drawing in the robot simulator. After visualizing the workspace in form of point clouds for each arm, the arm span and the volume of the shared workspace are calculated. This procedure is repeated for different numbers of random configurations. This is important for defining the reachability of the robotic system, e.g. the points in space it can reach.

Another important aspect of workspace analysis is finding the areas of reduced manipulability. The method for calculating manipulability in this paper was first described in [9]. The so-called Yoshikawa manipulability measure [10] describes how close the robot is to a singular configuration – e.g. how close it is to losing one of its degrees of freedom. The manipulability index is calculated as:

$$\boldsymbol{m} = \sqrt{\det(\boldsymbol{J}\boldsymbol{J}^{\boldsymbol{T}})} \tag{1}$$

where $\boldsymbol{J}$ is the Jacobian matrix of the robot arm at a certain configuration. This measure is proportional to the velocity ellipsoid at a given configuration. The velocity ellipsoid indicates the ability of the robot to move in each of the 3 translational and 3 rotational directions. If ***m*** equals zero it means that one degree of freedom is lost and that the robot is at a singular configuration. The bigger the index is, the more manipulability robot has in its current position. This measure is based only on the kinematics of the mechanism and does not take into account mass and inertia. Still, it's a good method for finding problem areas within the workspace that should be avoided in actual tasks.

The manipulability index is then calculated for all of the random configurations generated earlier. A cutoff index value is defined, below which a robot configuration is considered poorly manipulable. Percentages of such configurations for different number of random positions are calculated as well.

III. Simulation and Results

The system was modeled and the simulation was done in Matlab. Its' Robotics Toolbox package offers a number of ready-made functions, including the computing of forward, differential and inverse kinematics, trajectory planning, robot 3D animation, etc. After calculating the point cloud of reachable points, the workspace in body planes is shown in Matlab, while the visualization of the whole workspace and robot environment is done in CoppeliaSim robot simulator.

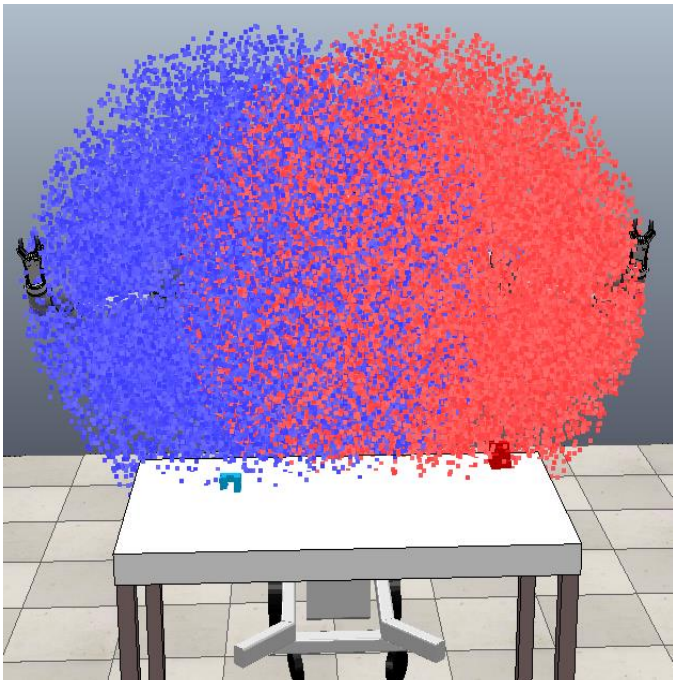


Fig. 4. Front view of the robot workspace in CoppeliaSim robot simulator.

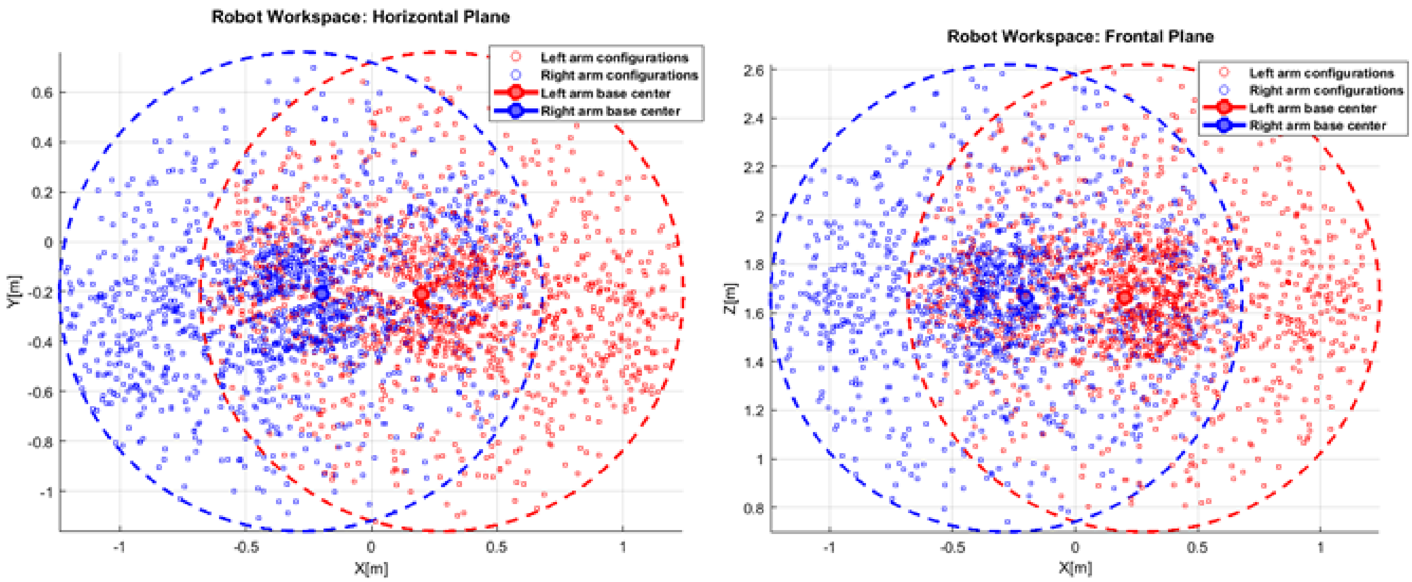


Fig. 3. Workspaces of the robot arms shown in horizontal (transverse) and frontal planes. The blue and red dashed lines mark the boundaries of the individual workspaces of the right and left robot arm workspaces, respectively.

Fig. 3 shows the robot workspace in horizontal and frontal body planes. The red points indicate the reachable points of the left arm while the blue ones are reachable for the right one. The individual arm workspaces are rounded with dashed lines so the shared workspace is easily identifiable.

Meanwhile, Fig. 4 shows the visualized workspace in robot simulator. The robot is in a slightly bent positon. A table with two objects (two parts) for an assemblage task is shown in front of it as an example of a typical task and typical robot environment. As in previous figure, the reachable points for the left and right arm are shown in red and blue, respectively.

So, the visualized workspaces show the reachable area for the robot arms and for the whole robotic system as well. Since each arm has a range of 950 mm the arm span of the whole robot structure is around 2300 mm. It has to be pointed out that UR5 arms themselves have a range of 850 mm. However, adding end effector equipment such as force sensors, wrist cameras and grippers extends that range. For this simulation only the grippers were added to the end effector resulting in a 100 mm larger range. The workspace volumes of the individual arms (the red and blue cloud points in Fig. 4) are around 3.6m^3. The volume of the shared workspace, visible in

Fig. 4 as the area where blue and red points intersect is around 1.12 m^3. It is also clear from Fig. 4 that the objects on the table are clearly within the reachable area.

From Fig. 3 it is clear that the individual workspaces of the robot arms do not just intersect in the area in front of the robot, but also in the back, as well as above and below the top of the torso. This information can be useful for specific tasks.

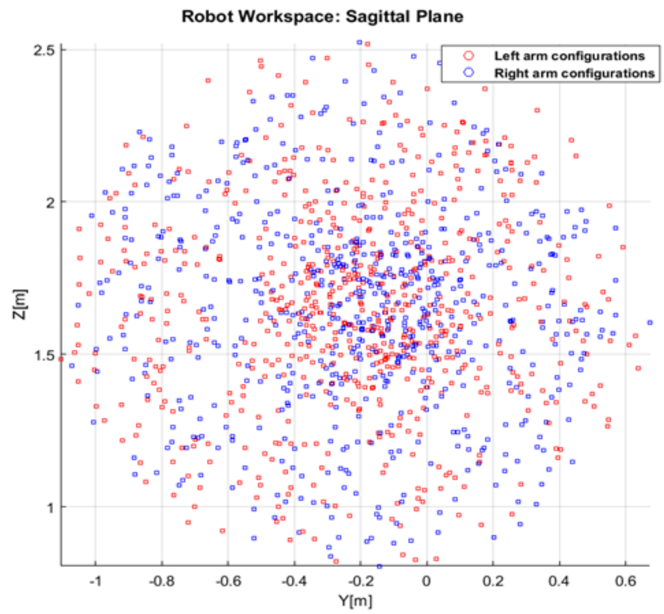


Fig. 5. Robot workspaces shown in sagittal plane.

Fig. 5 now shows the robot's sagittal plane. As its plane of symmetry, it contains similar number of configurations for both arms, equally and randomly distributed.

The robot structure itself is barely visible in Fig. 4. This is because of the relatively large number, 50000, of random configurations used for the simulation. Zooming into the point clouds in the robot simulator it can be seen that the points are equally distributed within the cloud (save for the areas in the torso which are not reachable). This is confirmed by Fig. 3 and Fig. 5 as well. These figures give the indication that all the areas within the workspace are equally reachable.

Manipulability analysis shows that that's not the case. Manipulability measure given with (1) is calculated for all the configurations used in mapping the workspaces. The key problem is identifying a key value of manipulability index, below which a configuration is considered poorly manipulable.

In [10] an index value of 10^{-5} is already considered quite poor, so a cutoff value of 10^{-4} is used to map problem areas within the workspace. Hence, the configurations with manipulability index of 10^{-4} or less are considered poorly manipulable (or "unmanipulable"). Configurations with the index between 10^{-4} and are 10^{-3} are relatively manipulable and those with the index bigger than 10^{-3} are not considered problematic.

Fig. 6. shows the robot system with the poorly manipulable points drawn. The points were calculated for each of the robot arms but were all drawn in the same color. The idea is to visualize them and identify problem areas to avoid for the whole robot in collaborative tasks.

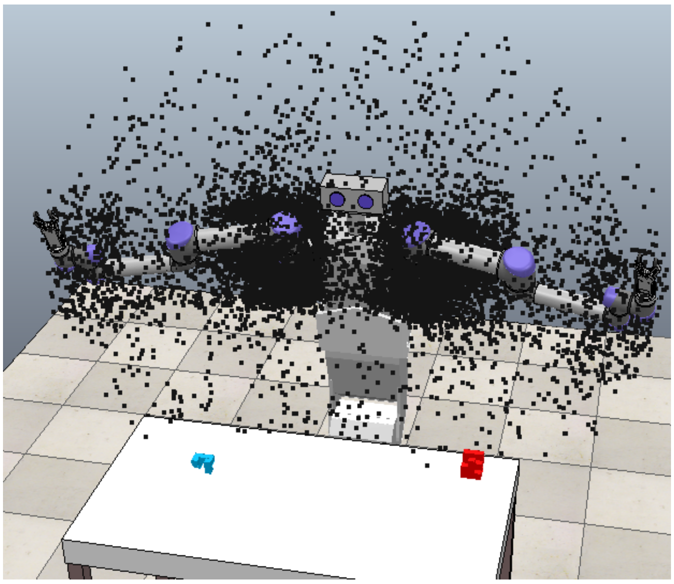


Fig. 6. Front view of the problem areas of the robot.

Clearly, there are clusters of problematic points around the bases of both arms and the first links as well as around the final links of the arms. On the other hand, there are no clusters in the area in front of the robot, closer to the table with the objects. Those are the least problematic areas which is important to know for practical tasks.

Fig. 6 shows only the points with the manipulability index equal to or less than 10^{-4}. If the relatively manipulable points were added too they would continue to cluster around the same locations. The areas which are in front of, behind, above and below the upper part of the torso, but which are a bit further from it, still remain the least affected by problematic points.

To corroborate this, and have a more detailed look into the structure of the problem areas it is good to again show the robot's body planes. This time they will be shown with unmanipulable configurations.

Fig. 7 shows the robots' horizontal and frontal planes. The problematic configurations are shown in the same color for both arms, to indicate problem areas for the whole robotic system, not just the individual arms. Again, it is evident that the least problematic areas are further away from the top of the torso in both directions.

These results again confirm the clustering of the problematic points around the top of the torso and to a lesser extent around the end effectors. The whole areas around the top part of the torso and the robot bases and first links form a clearly problematic region that should be avoided in tasks. However, this is not a practical problem for the robot. Since that region is very close to the location of the torso itself, many of those configurations are unreachable in any case, as the robot arms cannot be allowed to collide with the torso.

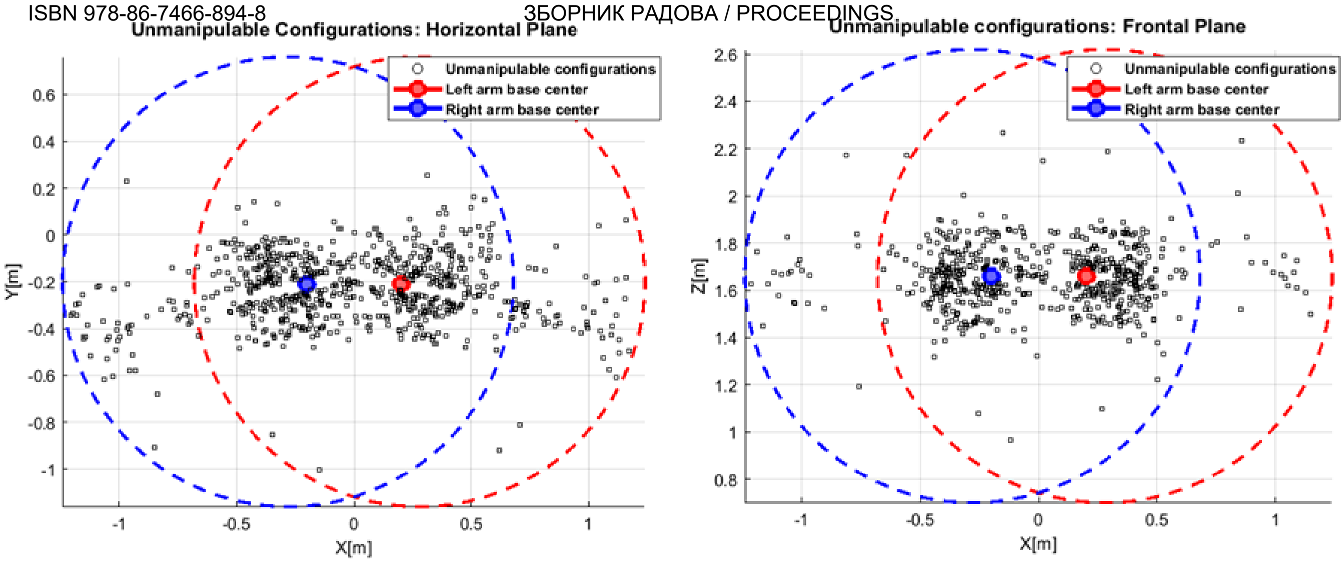


Fig. 7. Robot workspace with unmanipulable configurations shown in the horizontal and frontal plane of the robot's body.

It is interesting to analyze the number of poorly manipulable configurations, as well as the size of the manipulability index. The simulations were carried for a varying number of robot configurations, from 5000 to 50000 random positions.

The results are shown in Table I. They are shown for a single UR5 robot arm (the left one), since they are very similar to the other one.

TABLE I
MANIPULABILITY ANALYSIS

Number of configurations	Maximum manipulability index value	Minimum manipulability index value	Poorly manipulable configuration percentage
5000	0.1161	$8.472 \cdot 10^{-8}$	15.98
50000	0.1189	$1.124 \cdot 10^{-9}$	12.24
100000	0.1195	$5.710 \cdot 10^{-9}$	12.40
500000	0.1197	$1.338 \cdot 10^{-10}$	12.32

The final column shows the percentage of robot configurations with manipulability index less than 10^{-4}. It is the value below which a configuration is considered to have poor manipulability, as stated above. The percentage varies between approximately 16% for 5000 configurations to about 12.3% for a larger number of points.

The minimum value of the index is around 10^{-9}, which is pretty close to a singular configuration. The maximum value peaks at about 0.12, for the biggest number of robot configurations. While that indicates configurations with much more manipulability, the index is still not very high.

This shows that in practice, a seemingly large robot workspace can be significantly reduced by various limitations. Joint limits, self-collisions, singularities, and areas of reduced manipulability greatly impact its workspace and the way robot tasks can be defined and executed.

Another way of illustrating the manipulability in certain configurations is to show the robot with its velocity ellipsoids drawn. They are usually shown for the three translational components of the velocity. The volume of the ellipsoid is proportional to the manipulability at a given configuration. The longer the ellipsoid is along one of its axes, the more the robot can move in that direction and vice versa. If the configuration is singular the ellipsoid collapses into a planar ellipse as it loses one degree of freedom. If it loses another degree of freedom it can further collapse into a line.

The ellipsoids are shown here for three characteristic robot configurations. For the first one the robot arms are assembling an object directly in front of it, equally distanced from both arms. The second configuration is quite similar but the object is in front of the left arm base. The third is the initial configuration of the robot with its arms completely spread.

The results are shown in Fig. 8, 9 and 10. They are plotted in Matlab. The torso and the arms are shown as a kinematic chain and the ellipsoids are plotted at the arms' end effectors. The colors are again red for the left arm and blue for the right one.

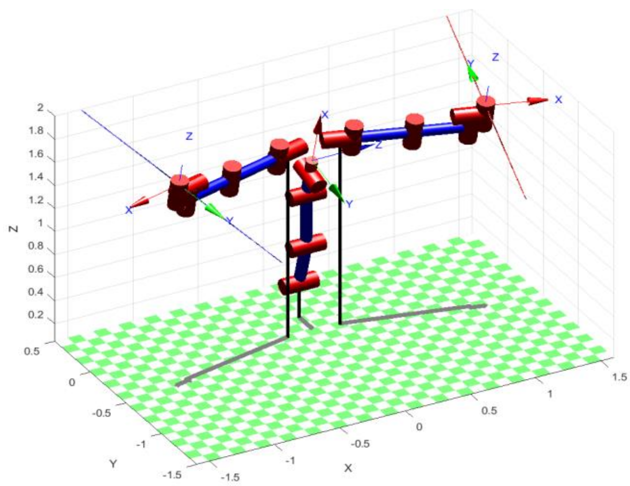


Fig. 8. Velocity ellipsoids in initial position.

The initial position is shown first and it is immediately clear that it is a singular configuration for the arms. Only lines are plotted meaning the robot loses two degrees of freedom and can only move in one translational direction.

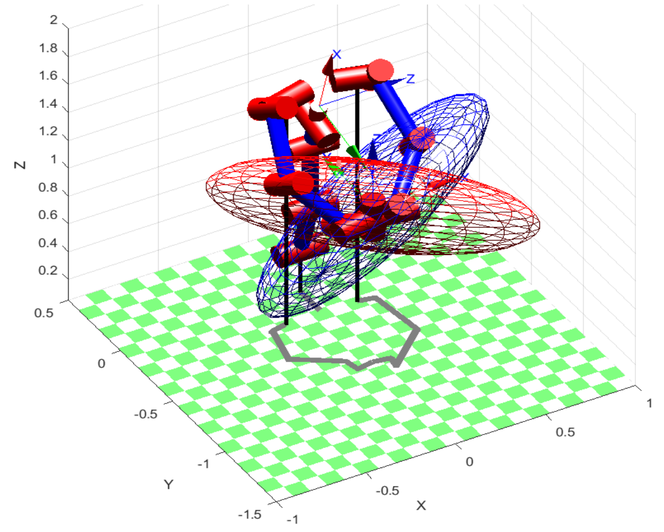


Fig. 9. Velocity ellipsoids in an assembly task in front of the torso

Fig. 9 shows the assembly position in front of the torso. The object is at an equal distance from both arms, and the ellipsoids are clearly much bigger, and equal to each other, indicating good manipulability.

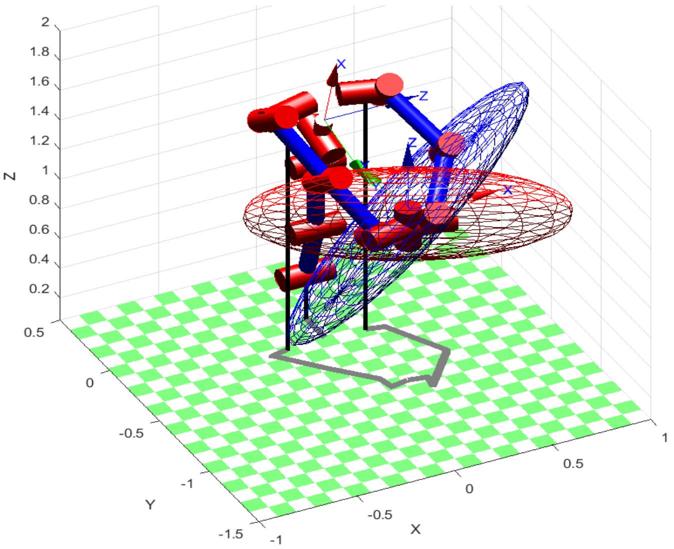


Fig. 10. Velocity ellipsoids in an assembly task in front of left arm base

Fig. 10 shows a similar task, this time in front of left arm basis. The shape of the left arm ellipsoid (red) remains similar while the right one is longer in one direction but quite shorter in another, which indicates a decreased manipulability for the right arm, which is expected.

IV. Conclusion

Simulation results have given a good overview of the workspace of a bi-manual robotic system. They have shown a large reachable area of the whole robot as well as a considerable shared workspace. Defining this workspace is very important as one of the early steps in developing a custom bi-manual robotic system. This information is essential for configuring the robot's environments and for defining its various future tasks.

Manipulability analysis has identified problem areas within the reachable workspace as well as areas of good manipulability. This is important for defining the way the robot will perform its tasks and its optimal trajectories.

While a good indicator of the robot's manipulability, this analysis is not perfect. It only takes into account the kinematic properties of the robot, and it does not mean that all of the unmanipulable or relatively manipulable configurations will present a problem for the physical robot. Still, it does give a good overview of the areas that should be avoided.

This paper presented a starting work in the analysis of a custom bi-manual robotic system. A more detailed manipulability analysis, that takes into account the masses and inertias of the mechanism could be performed, and compared with previous results. The authors' future work will also consider various other aspects of bi-manual robotic systems such as different control strategies, trajectory planning, etc.

V. Acknowledgements

The results presented in the paper are obtained in the scope of the research projects: "Development and Experimental Performance Verification of Mobile Dual-Arms Robot for Collaborative Work with Humans", Science an Development Programme - Joint Funding of R&D Projects of the Republic of Serbia and the People's Republic of China, contract no. 401-00-00589/2018-09, 2018-2021 and national R&D project no. TR-35003, both supported by the Ministry of education, science and technology development of Republic Serbia.

References

[1] C. Smith, Y. Karayiannidis, L. Nalpantidis, X. Gratal, P. Qi, D. V. Dimarogonas, and D. Kragic, "Dual arm manipulation - a survey," *Robot. Auton. Syst.*, vol. 60, no. 10, pp. 1340–1353, 2012.

[2] M. Fuchs et al., "Rollin' Justin - Design considerations and realization of a mobile platform for a humanoid upper body," 2009 IEEE International Conference on Robotics and Automation, 2009, pp. 4131-4137, doi: 10.1109/ROBOT.2009.5152464.

[3] Z. Pilat, W. Klimasara, M. Pachuta, M. Słowikowski, "Some New Robotization Problems Related to the Introduction of Collaborative Robots into Industrial Practice" *J. Autom. Mob. Robot. Intell. Syst*. vol. 13. no. 4, pp. 91-97, 2020.

[4] R.L. Williams II, "Baxter Humanoid Robot Kinematics", Internet Publication, April 2017, https://www.ohio.edu/mechanical-faculty/williams/html/pdf/BaxterKinematics.pdf.

[5] ABB IRB 1400 Yumi Technical Data, May 2021, https://new.abb.com/products/robotics/collaborative-robots/irb-14000-yumi/irb-14000-yumi-data

[6] EPSON Technical Data, May 2021, https://www.epson.eu/robots/dual-arm

[7] A. Rodić, J. Šumarac, I. Stevanović, M. Jovanović, "Cloud-Enabled Bi-manual Collaborative Robot with Enhanced Versatility for Customized Production", Proceedings of the 30th International Conference on Robotics in Alpe-Adria-Danube Region, Futuroscope-Poitiers, France, pp. 240-249, June, 2021.

[8] Universal Robots Technical Data, May 2021, https://www.universal-robots.com/download/manuals-cb-series/service/service-manual-ur5-cb3-english/

[9] P. Corke, *Robotics, Vision and Control*, 2nd ed., Berlin, Germany, Springer, 2017.

[10] T. Yoshikawa, "Manipulability of Robotic Mechanisms", *Int. J. Rob. Res.*, vol. 4, no. 2, pp 3-9, 1985.

Interface development dedicated to connecting CAD tools for 3D modeling of complex objects and industrial robot's controller

Uroš Ilić, Jovan Šumarac, Ilija Stevanović, Aleksandar Rodić

Abstract—**This article covers the development of the interface dedicated to converting CAD models into some universal computer readable form – Excel spreadsheet. Main intention is transcription of complex models' geometry and coloring to numerical database. First the article will describe the process of saving CAD models in compatible form and its conversion to Microsoft Excel spreadsheets via a set of MATLAB's designated scripts and functions. Main focus is forming a set of significant points meant to guide the industrial robot's end effector to outline (with a laser pointer) the surface of the object that's in contact with the work desk. At the end, the given trajectory will be shown in MATLAB's plotter and correspondent Excel table.**

***Index Terms*—CAD; MATLAB; Microsoft Excel; Path planning; Industrial robots**

I. Introduction

MASS production as a leading example of XX century rapid development is slowly but consistently being replaced by mass customization. Its elementary feature is the use of flexible computer-aided manufacturing systems to produce custom output. Such systems combine the low unit costs of mass production processes with the flexibility of individual customization.

Even though integrating robots to manufacturing and assembly processes started long time ago, new problem opposes: How to program the robot in such way so production lines become more flexible in terms of fulfilling customer demands? For example, if it happens that there is a specific requirement on only one product, it is necessary to program the entire workstation in accordance with the requirement. After performing the task, it needs to be reprogrammed for the prior purpose. All this work is forming expenses, both the financial ones and also in form of human resources, i.e., the engagement of engineers and/or programmers.

With the advent of factory digitalization and Industry 4.0, the whole world is moving towards factories with no or minimal number of workers. For this purpose, we strive for remote control of production systems, i.e., robots, both on production tasks and on assembly lines. With that in mind, it is necessary to develop such system where it is possible to control individual robots (manipulators). One of the conceptual solutions is forming a database within the cloud, which could be accessed simultaneously by the robot, which performs the assembly, and a programmer, who would work on updating the database itself.

II. Objectives Of CAD Systems And Robotic Controller Integration

Each of the designed parts will, firstly, be modeled in Solidworks® environment, and then exported with the appropriate extension. Such files are imported into MATLAB and, afterwards, exported to Microsoft Excel spreadsheets. This type of data storing is found convenient for later use by numerous programming languages such as C, C++, Python or MATLAB itself once again. For each of the imported models in MATLAB, a set of designated points in space with TCP's approach vector will be formed.

The final goal is to generate the trajectory of TCP with the given information, and thus solve the inverse kinematic problem of the manipulator. In other words, it is necessary to form a sequence of matrices of significant points' coordinates along with approach vector's x-y-z components, which will unambiguously determine the position of the robot's segments when moving, i.e., performing a certain technological operation.

All gathered information will be integrated into the database and the robot will, based on the data it receives from the camera, with the help of artificial intelligence, define which parts are in front of it, and in which position and orientation. Based on all of that, the robotic system will choose the appropriate technological operation of assembly or manufacturing process. Finally, with the given information, the robot can initialize, first the formation of the trajectory, and afterwards its realization.

Uroš Ilić is with the Faculty of Mechanical Engineering, University of Belgrade, 16 Kraljice Marije, 11020 Belgrade, Serbia (e-mail: ilicuros01@gmail.com)

Jovan Šumarac is with the Institute Mihajlo Pupin, University of Belgrade, 15 Volgina, 11020 Belgrade, Serbia (e-mail: jovan.sumarac@pupin.rs)

Ilija Stevanović is with the Institute Mihajlo Pupin, University of Belgrade, 15 Volgina, 11020 Belgrade, Serbia (e-mail: ilija.stevanovic@pupin.rs

Aleksandar Rodić is with the Institute Mihajlo Pupin, University of Belgrade, 15 Volgina, 11020 Belgrade, Serbia (e-mail: aleksandar.rodic@pupin.rs)

III. Using Solidworks® Software For Designing Complex Objects

Solidworks® is a solid modeling computer-aided design (CAD) program that's suitable for designing numerous complex objects. In this particular case various mechanical components are modeled and used for development of the interface. Some of these objects are shown on Figure 1.

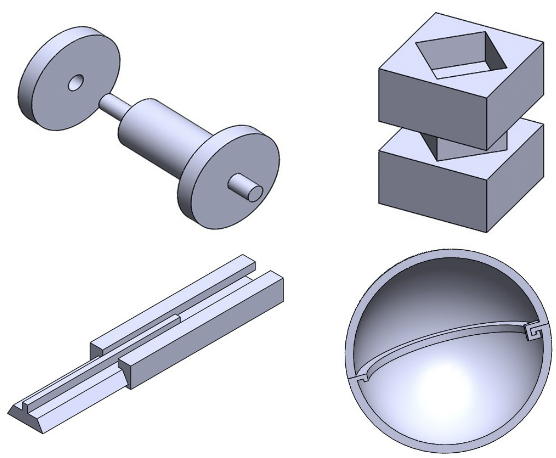


Fig. 1. Representation of typical complex objects found in the machine industry. These objects are chosen as they represent various kinds of assemblies, depending on interrelation of the parts.

The selected assemblies are some of the most common types in the machine industry and assembly processes. They are listed gradationally, so that simpler assemblies are described first. The simplicity of the assembly is observed from the aspect of the number of possible mutual positions that the parts can or must take to form an assembly.

A. Axle with bearings

This is one of the simplest mounting assemblies. In layman's terms, bearings can be found in infinitely many positions with the axle. This statement is justified by the axial symmetry of the parts that enter the assembly. It is necessary to achieve the appropriate positions immediately before the assembly process and perform a simple translation in order to form an assembly. Their model is shown in Fig. 1 in upper-left corner. Even though bearings need significant force to be mounted, this assembly is only used in the field of robot kinematics.

B. Simple puzzle

This form represents a more complex assembly than the previous one since only 4 mutual positions in the assembly are possible. They are achieved by a 90° rotation around the mounting axis. If the parts are not marked, the observer will not even be able to distinguish the different positions. Even though this type of the simplest puzzle is not typically found in the industry it's just a mere representation of assembly of prismatic shapes. It is shown in Fig. 1 in upper-right corner.

C. Linear guide rail

Classical example of a linear bearing, that is actually an assembly that can only be formed in a certain way. The guide and the slider must be brought to a precisely determined position and orientation in order for the assembly to be formed by a simple translation or inserting the slider into the guide. This assembly is shown on Fig. 1, bottom-left corner.

D. Spherical vessel

Unlike the other parts, it is not given in some of the standard views (isometry, trimetry ...). The given view (Fig. 1, bottom-right corner) is selected to show the grooves for closing and opening. This type of assembly is one of the most complex assemblies that can be found on the assembly line. It is necessary to bring both hemispheres to a position in which the tongue of the upper part coincides with the slit of the lower part. After connecting the hemispheres, it is necessary to rotate them by a certain angle and thus achieve the formation of the assembly, i.e., sealing the spherical vessel.

E. 3D puzzle pyramid

This type of assembly was taken into consideration as an example of the most complex technological assembly operation. The pyramid consists of 5 parts. It is necessary to assemble them in a specific order. An additional aggravating circumstance is the difficulty of distinguishing the parts of the pyramid itself. For this purpose, the outer sides of the parts are painted with various colors, while the inner sides are painted with the characteristic color of wood. The idea is to assemble the pyramid by a two-handed manipulator in some of the future works. Both assembled and disassembled pyramid is shown in Figure 2, respectively.

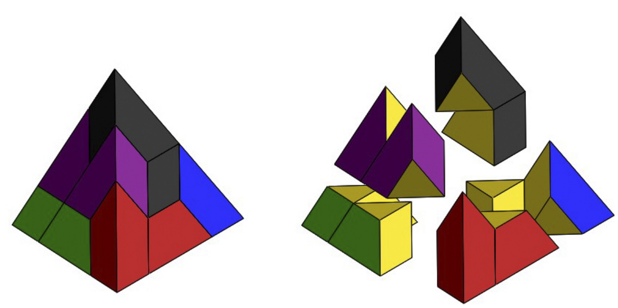


Fig. 2 – Comparison of assembled and disassembled pyramid. The complexity of parts' interrelationships can be seen on the figure.

IV. Surface Triangulation, Data Export And Objects' Database Creation

In order for the parts and its CAD models to be readable by the robot's control unit, it is necessary that the geometry data of those parts is stored in numerical format. For this purpose, Microsoft Excel spreadsheet is chosen. Apart from being compatible with numerous programming languages, it is also easy to be reviewed by the user of the interface.

Of course, direct conversion from CAD models to Excel spreadsheets is not possible. Therefore, the MATLAB software package will be used, where the procedures for the indirect conversion of parts from the Solidworks® environment into an Excel spreadsheet will be written. In order for all this to be feasible, it is necessary to export the CAD model in the appropriate format. Figure 3 represents the flowchart of this method.

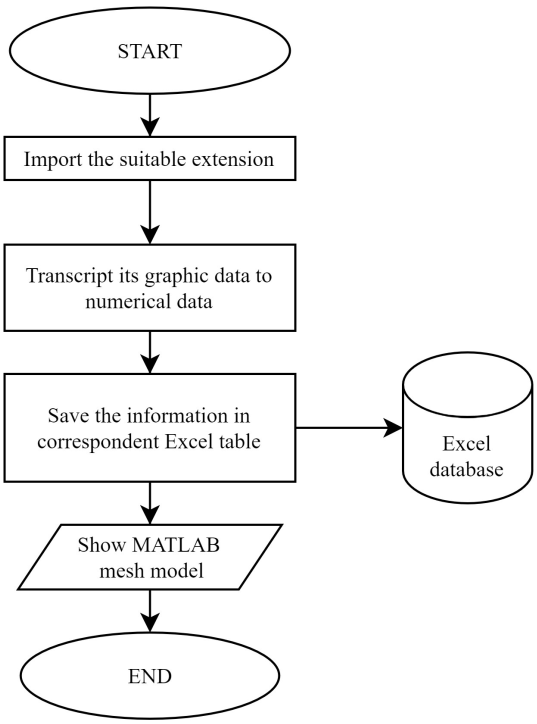


Fig. 3 – Flowchart representing the process of transferring the CAD model to the designated Excel spreadsheet with a mesh model's representation in MATLAB plotter

Of all the possible extensions that Solidworks® supports, only two solutions are significant - STL format (.stl) and VRML format (.wrl). Although the selection of each of them entails certain advantages and disadvantages, that are presented in Table 1.

TABLE I
COMPARISON OF CHOSEN EXTENSIONS' CHARACTERISTICS

Format (extension)	Advantages	Disadvantages
VRML (.wrl)	Ability to work with colors and their RGB vectors	Poorly researched topic; Lack of support on MATLAB forums
STL (.stl)	Relatively easy import to MATLAB; Very strong support within the MATLAB community	Able to store only geometry data, but not coloring of the faces

A. STL files

The abbreviation STL (Standard Triangle Language) represents a format suitable for 3D printing, i.e., stereolithographic methods in general. The format is characterized by triangulation of surfaces, i.e., division of surfaces into triangles. Figure 4 shows the process of forming a STL file.

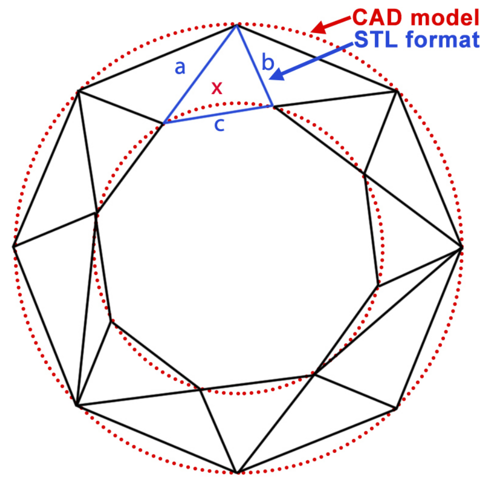


Fig. 4 – From the amorphous structure of the torus shape represented with dotted red line in the picture, a multitude of triangles is formed (one of which is marked in blue).

If it is amorphous, i.e., the more complex the surface, the greater the number of triangles in the network and vice versa. It is quite certain that simple geometric surfaces such as rectangles or squares will be divided into only two triangles, while circles and elliptical shapes will be divided into a finitely large number of triangles. Every STL file is characterized with connectivity matrix, coordinates of every point and each triangle's normal vector. The connectivity matrix has dimensions 3xn, where n represents number of triangles. This matrix defines what points form what triangle, respectively. Each triangle has its own corresponding normal vector represented with a unit vector.

B. VRML files

The VRML (Virtual Reality Modeling Language) format was created at the end of the last century with the intention of displaying three-dimensional vector graphics objects. The main emphasis was on the implementation of this format in internet communication, especially web presentations. These types of files (although their extension is *.wrl) are actually textually formatted and very easy to read by humans. Of course, by reading only a text file, one cannot create the body image that the file describes. Therefore, MATLAB will be used, which will read certain segments of the file, convert them into numerical values and form from them, first a mesh model, and then an Excel spreadsheet.

As mentioned in Table 1, VRML files have the ability to store information about faces' coloring. Information about coloring is stored in similar way like those for normal vectors. For each generated triangle a three-dimensional vector is formed that describes primary colors in additive color synthesis. Each component represents the amounts of red, green and blue, respectively.

C. Creating the database

After successfully exporting the models in suitable form, it's needed to import them in the MATLAB's environment. After loading the file's full name, the designated function makes the spreadsheet with the same name as the STL/VRML file. Given spreadsheet is formatted in such manner (merging cells, naming the titles, etc.) that is appropriate for following code to fill in the information about triangulation (connectivity matrix), coordinates, normal vectors and, regarding the VRML file, the coloring too.

When reading of the chosen format is done, designated script stores significant values in appropriate matrices and starts filling in the Excel spreadsheet. If chosen format is STL, the script uses modified version of downloaded function stl_read(), made by MATLAB community. On the other side, if it's needed to convert the VRML file, the script calls designated function wrl_read(), that, firstly convert's the VRML file to text file and extracts desired strings of text that contains information that's needed to fill in the Excel spreadsheet. These strings are converted into numerical form and saved as matrices, that, as in previous case, fill in the database.

In Figure 5 a typical Excel spreadsheet with CAD model's info is represented. The spreadsheet is cut in half over the J-column in order to fit the formatting of this article.

	A	B	C	D	E	F	G	H	I	J
1	Triangles					Vertices (coordinates)				
2	Number	Vertices (number of vertex)				Number	X	Y	Z	
3	1	7	2	9		1	0	-84	0	
4	2	9	2	5		2	0	0	0	
5	3	4	6	9		3	8.083	-70	22.862	
6	4	5	2	4		4	8.083	-42	22.862	
7	5	4	2	1		5	8.083	-14	22.862	
8	6	4	1	3		6	24.249	-42	0	
9	7	2	7	1		7	24.249	-14	0	
10	8	1	7	6		8	32.332	-56	22.862	
11	9	1	6	10		9	32.332	-28	22.862	
12	10	9	6	7		10	48.497	-56	0	

	J	K	L	M	N	O	P	Q	R	S
1		Normal vectors (unit vectors)					Triangles' colors			
2		Number	X	Y	Z		Number	R	G	B
3		1	0.471405	0.816497	0.333333		1	1	0.784314	0
4		2	0.471405	0.816497	0.333333		2	1	0.784314	0
5		3	0.471405	-0.816497	0.333333		3	1	0.784314	0
6		4	-0.942809	0	0.333333		4	1	0.784314	0
7		5	-0.942809	0	0.333333		5	1	0.784314	0
8		6	-0.942809	0	0.333333		6	1	0.784314	0
9		7	0	0	-1		7	1	0.784314	0
10		8	0	0	-1		8	1	0.784314	0
11		9	0	0	-1		9	1	0.784314	0
12		10	0.942809	0	-0.333333		10	1	0.784314	0

Fig. 5 – The appearance of the Excel spreadsheet that originated from model's VRML file. This screenshot shows data only for first 10 triangles, even though the number of triangles can be measured in thousands for amorphous surfaces.

Here, a user can see from what vertices the triangles are format, e.g., the vertices numerated as 7,2 and 9 form the first triangle. Their coordinates in local coordinate system can be read in the neighboring table. Components of the normal unit vector for the first triangle can be read from cells L3 to N3 and its RGB coloring vector can be found in cells Q3:S3. Note that this spreadsheet came from conversion of the VRML file, since it has information about the coloring. The spreadsheet formed from STL files is practically the same without the last (fourth) table describing the colors.

Note that the tables always start from the same cell, no matter what model is being converted. Only change is number of rows in the Excel file. This feature enables making the universal MATLAB function for reading Excel spreadsheets and storing the values in corresponding matrices.

V. GENERATING MESH MODELS FROM 3D DATABASE

As a verification of the stored data, it's needed to read the spreadsheet and visualize that numerical information with a mesh model. If and only if the mesh model made by MATLAB plotter is same as initial CAD model, the code can be verified.

Firstly, the code calls the Excel file with desired name and starts reading the data from the tables. For every table read, it stores the values in corresponding matrix. After that, an already built-in functions form a three-dimensional mesh from the connectivity list (first table in the spreadsheet) and corresponding coordinates of given vertices (second table in the spreadsheet). Afterwards, each normal unit vector is drawn with the origin in triangle's centroid. If we're working with the STL files, all triangles will be painted with the same color. On the other side, if the user is converting the VRML file, each triangle will be colored with their correspondent RGB vector. Figure 6 shows the CAD model of an axle on the left and the result of reading the spreadsheet on the right. This is the result of converting the STL files.

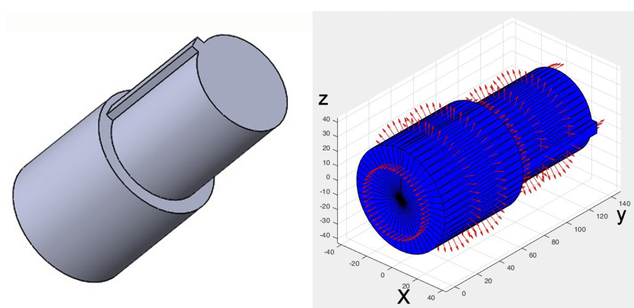


Fig. 6 – The result of the conversion of the CAD model to Excel spreadsheet and back again to MATLAB's mesh model. Coloring of the mesh model can be changed to any color, but whole model must be colored uniformly

On the other side, if we chose to export the CAD model as a VRML file, it will store its original coloring and represent it in its original colors in MATLAB plotter. Figure 7 shows a mesh model with surface vectors of 3D puzzle pyramid's red part in its original coloring. As it was said earlier the parts' outer surface is colored differently, but their inner sides depict wood's natural coloring.

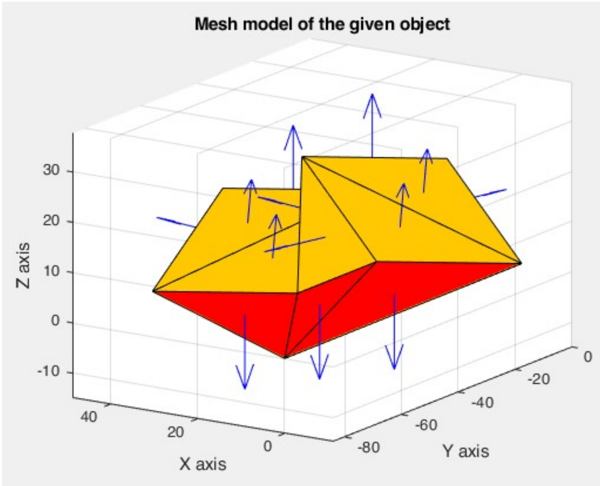


Fig. 7 – 3D puzzle pyramid's corresponding mesh model of the red part shown in MATLAB plotter. Outer face is colored with red and inner ones are with color of wood, just like we colored them in Solidworks®. Blue arrows pointing outwards represent normal unit vectors of each triangle.

VI. Defining The Contour Of The Object And Corresponding Significant Points

The procedure for defining the path that will enable the TCP to outline the surface lying on the desk is given on the figure 7 in form of the flow chart. This flowchart represents the second part of the whole process – the extraction of the numerical data of modeled parts.

The first step is to acquire the information of the object on the work desk. The camera will recognize the object lying on the desk and its position and orientation. Let's say that camera recognized the object on Figure 7. With the help of artificial intelligence, desired object's information will be found in the database. With the knowledge of position of the object in global, i.e., robot's coordinate system, object's local coordinate system will be transformed. Hence the vertices in contact with the desk will change. Therefore, we'll know the exact vertices lying on the desk, and their coordinates in global coordinate system.

The problem occurs if the shape of the face is oval or circle-shaped. Since the triangulation transforms the circle into a polygon, the center of the circle becomes the vertex that' included in all triangles. Even though it has its z-coordinate equal to zero, we must eliminate it too in order to form a closed polygon.

With the points extracted from the Excel spreadsheet, the 2D polygon of the surface is made. On Figure 8 the polygon is drawn with red lines with normal vectors pointing outside. The face represented on the Figure is actually the outer side of the puzzle's part – the red side.

When the polygon representing the contour of the object lying on the work desk is constructed, we advance to forming a trajectory of the tip of the end-effector of the robot. As said earlier, the laser pointer will be attached to the end-effector that will outline the contour of the object. In order to do such a thing, we must make another polygon with desired offset from the edges of the initial polygon.

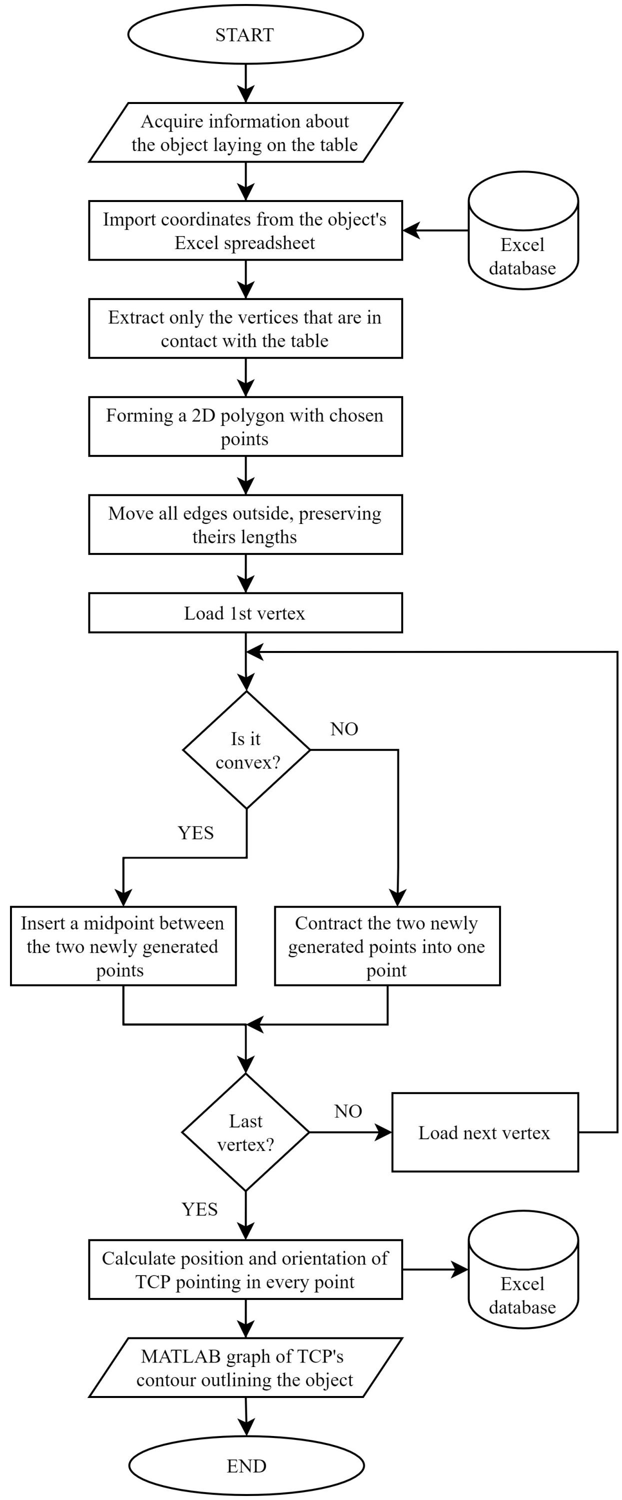


Fig. 7 – A flowchart depicting the process of the creation of significant points of TCP's path outlining the lying surface of the given object

The newly formed polygon is actually a modified enlargement of the surface's contour with rounded edges. It is constructed in a *for*-loop that goes over every point and checks if it's in a convex or concave vertex. The whole process of creating the rounded enlarged polygon from the initial one can be described with the following steps:

- First, every edge is shifted outwards for desired distance. Now, instead of closed polygon, we have a set of disconnected lines, but with preserved lengths.
- For every vertex of the contour, algorithm checks whether they're convex or concave
- If they are convex, a new middle point is added, so that end effector outlines the convex vertex in a radius
- If, on the other hand, the vertex is concave, two significant points are replaced with one and therefore radius is made again in order to avoid collision.

After this part of the algorithm completes the whole contour and the last vertex is formed, the potential trajectory of the TCP is plotted in 2D diagram, as represented in Figure 8.

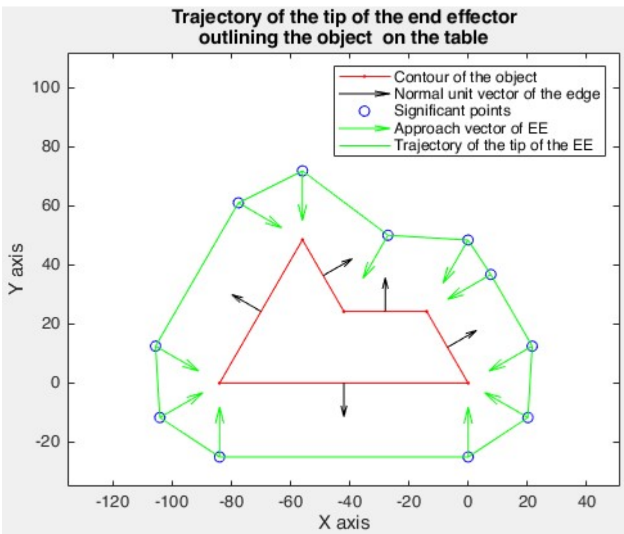


Fig. 8 – Newly generated trajectory of the end effector. It can be seen that one of the vertices is concave, hence a collapsing of neighboring points occurred. It can also be seen that the edges have been shifter outwards in the direction of normal unit vector.

In every, newly generated, significant point, the end effector's approach vector is computed. Approach vector will always have direction from the significant point to the correspondent point within the contour. Every point's position and approach vector's components are then stored in an Excel spreadsheet, as shown on Figure 9.

Points	TCP Position			Approach vecotor		
	X	Y	Z	X	Y	Z
1	0.00000	-25.00000	43.30127	0.00000	0.50000	0.86603
2	20.20032	-11.66273	43.30127	-0.43301	0.25000	0.86603
3	21.65070	12.49989	43.30127	-0.43301	-0.25000	0.86603
4	7.65070	36.74889	43.30127	-0.43301	-0.25000	0.86603
5	-0.04962	48.41162	43.30127	-0.25000	-0.43301	0.86603
6	-27.11194	50.03582	43.30127	-0.25000	-0.43301	0.86603
7	-56.00010	71.82244	43.30127	0.00000	-0.50000	0.86603
8	-77.65059	60.99708	43.30127	0.43301	-0.25000	0.86603
9	-105.65059	12.50008	43.30127	0.43301	-0.25000	0.86603
10	-104.20031	-11.66261	43.30127	0.43301	0.25000	0.86603
11	-84.00000	-25.00000	43.30127	0.00000	0.50000	0.86603

Fig. 9 – The look of the Excel spreadsheet containing the information about TCP's position (left part) and end-effector's approach vector outlining the object's surface on the table. The Z coordinates are the function of the desired angle of the laser's beam.

VII. Conclusion

Although the given example is a simple problem, it represents the new way of transcribing and eventually storing the numerical representation of complex geometry objects. This way of storing the data (in Excel spreadsheets) is chosen because it is the best way for a human eye to examine the data. This data can be stored in any file that is suitable for reading as long as the programmer knows how to store/look/extract the information.

It is clear that this process doesn't give homogeneous transformational matrices for every significant point. The idea is to store information in that way that can be easily retrieved by robot's controller or any other computing unit meant for simulation or robot control.

Acknowledgment

The results presented in the paper are obtained in the scope of the research projects: „Development and Experimental Performance Verification of Mobile Dual-Arms Robot for Collaborative Work with Humans", Science and Development Program – Joint Funding of R&D Projects of the Republic of Serbia and the People's Republic of China, contract no. 401-00-00589/2018-09, 2018-2021 and national R&D project no. TR-35003, both supported by the Ministry of education, science and technology development of Republic Serbia.

References

[1] Peddie, Jon, *The history of visual magic in computers: how beautiful images are made in CAD, 3D, VR and AR*, 1st ed, London, Springer,2013

[2] Tseng, M.M.; Jiao, J. (2001) "Mass Customization", in: *Handbook of Industrial Engineering, Technology and Operation Management* New York, NY: Wiley, 2001, 978-0-471-33057-8

[3] J. J. Uicker, G. R. Pennock, and J. E. Shigley, *Theory of Machines and Mechanisms*, New York, Oxford University Press, 2003

[4] Paul Festa and John Borland, "Is a 3D web more than just empty promises?". CNET News.com, May 19, 2005.

[5] P. Corke, *Robotics vision and control. Fundamental algorithms in MATLAB®,* 3rd ed, London, Springer, 2011

A Mobile Robot Visual Perception System based on Deep Learning Approach

Aleksandar Jokić, Lazar Đokić, Milica Petrović and Zoran Miljković

***Abstract*—In this paper, we present the novel mobile robot perception system based on a deep learning framework. The hardware subsystem consists of an Nvidia Jetson Nano development board integrated with two parallelly positioned Basler daA1600-60uc cameras, while the software subsystem is based on the convolutional neural networks utilized for semantic segmentation of the environment scene. A Fully Convolutional neural Network (FCN) based on the ResNet18 backbone architecture is utilized to provide accurate information about machine tool models and background position in the image. FCN model is trained on our custom-developed dataset of a laboratory model of manufacturing environment and implemented on mobile robot RAICO (Robot with Artificial Intelligence based COgnition).**

***Index Terms*—Deep learning; Perception System; Mobile robot; Semantic Segmentation.**

I. Introduction

Modern mobile robot sensors (e.g., cameras or lidars) provide a rich amount of data about the current state of the environment. However, the way the data is interpreted and transferred into useful information has been an active area of research in the last two decades. Deep learning, or more precisely, Convolutional Neural Networks (CNNs), represent one of the most promising methodologies that can enable mobile robots to understand and interact with their environment in a more sophisticated manner [1]. The main disadvantage of CNNs is the requirement for a substantial amount of computation power for real-time implementation. Fortunately, several modern single-board computers or hardware accelerators provide enough computing power to deploy low-weight CNNs for real-time implementation.

The authors [2] developed the visual perception system utilized for navigation in the indoor environment based on CNNs for scene classification. They deployed shallow CNN to achieve real-time mobile robot navigation within the environment in both dynamic and static conditions. The development of the CNN model capable of determining the 3D physical properties of objects in the scene is presented in [3]. The authors propose using the learned properties to predict the outcome of the dynamic events in the environment. This type of perception system can be beneficial for mobile robots employed in highly dynamic environments. The authors of [4] developed a complex cleaning mobile robot perception system with two submodules, one based on Bayesian filtering of data from 2D lidar, 3D lidar, and RGB-D camera used for human detection and tracking, and the second one for obstacle and dirt detection based on two RGB-D cameras and 3D lidar sensor. In [5], the authors developed a perception system based on a monocular camera for data acquisition and SURF point feature extraction method integrated with neural extended Kalman filter for simultaneous localization and mapping of mobile robot's position and orientation. Another promising methodology for developing perception systems for mobile robots is a cooperative perception [6], where the perception of one autonomous agent depends on the perception of other nearby agents. Camera information obtained by one mobile robot (agent) can be propagated to the others if their relative pose is estimated well. Moreover, the authors implemented the camera models that share information regarding pedestrian detection. The mobile robot developed with the aim to be applied for manufacturing purposes was proposed in [7]. Safe indoor navigation is provided by the perception system based on a 2D camera, 3D camera, laser scanner, ultrasonic sensors, and internal measurement unit. On the one side, safe navigation is provided by a laser scanner and ultrasonic sensors, while a 3D camera is used for the correction step in the Kalman filter state estimation. Moreover, the 2D camera is utilized for adaptive path planning by detecting lines on the ground.

This work presents the novel semantic segmentation-based mobile robot perception system implemented on our own developed mobile robot RAICO. The modified ResNet18 backbone architecture is integrated with RAICO's sensory subsystem, which contains two parallelly positioned Basler daA1600-60uc cameras. The proposed perception system provides RAICO with the ability to safely navigate the laboratory model of the manufacturing environment by knowing the position of machine tools in the image plane.

Aleksandar Jokić, PhD student, University of Belgrade - Faculty of Mechanical Engineering, Department of Production Engineering, Laboratory for industrial robotics and artificial intelligence (ROBOTICS&AI), Kraljice Marije 16, 11120 Belgrade 35, The Republic of Serbia (ajokic@mas.bg.ac.rs).

Lazar Đokić, PhD student, University of Belgrade - Faculty of Mechanical Engineering, Department of Production Engineering, Laboratory for industrial robotics and artificial intelligence (ROBOTICS&AI), Kraljice Marije 16, 11120 Belgrade 35, The Republic of Serbia (ldjokic@mas.bg.ac.rs).

Dr. Milica Petrović, Assistant Professor, University of Belgrade - Faculty of Mechanical Engineering, Department of Production Engineering, Laboratory for industrial robotics and artificial intelligence (ROBOTICS&AI), Kraljice Marije 16, 11120 Belgrade 35, The Republic of Serbia (mmpetrovic@mas.bg.ac.rs).

Dr. Zoran Miljković, Full Professor, University of Belgrade - Faculty of Mechanical Engineering, Department of Production Engineering, Laboratory for industrial robotics and artificial intelligence (ROBOTICS&AI), Kraljice Marije 16, 11120 Belgrade 35, The Republic of Serbia (zmiljkovic@mas.bg.ac.rs).

The paper is structured as follows. Section two is devoted to a thorough explanation of the considered CNN model and its training procedure. The third Section describes the experimental results and perception system evaluation while concluding remarks are presented in the fourth Section.

II. Mobile Robot Perception System

The mobile robot perception system is developed with two parallelly mounted Basler daA1600-60uc cameras facing downwards with an inclination angle of 30° and a baseline of 12.5 cm. The combination of mentioned cameras with *Evetar M118B0418W* lenses (4 mm focal length) provides a large angle-of-view of the scene, approximately W×H=84°×68°. The cameras are connected via USB3.0 to the Jetson Nano development board for image acquisition. The whole perception system is positioned on the top of the mobile robot RAICO (Fig. 1).

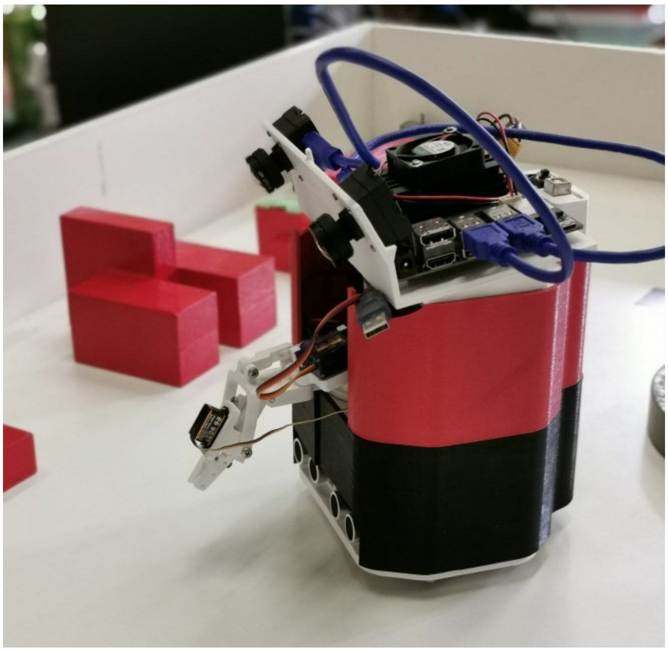


Fig. 1. Mobile robot RAICO with assembled perception system

The main component of the perception system is the Fully Convolutional neural Network (FCN) used for semantic segmentation. To enable the processing in near real-time, the selected backbone is ResNet18 based network. The network is trained by utilizing our custom-developed dataset for semantic segmentation. Image data is acquired within the laboratory model of a manufacturing environment, while the segmentation masks are hand-labeled. The dataset consists of densely labeled images with four machine tool classes and the background class. The dataset contains 125 images divided for training and testing in the 80/20 ratio. The sample of the dataset is shown in Fig. 2.

Hard data augmentation is carried out on the images used for training to improve neural network generalization. Mobile robot RAICO uses real-world noise-prone cameras that can significantly impact the accuracy of neural networks [8]. Therefore, the first augmentation is performed with Gaussian noise added to the images. Two Gaussian noise levels have been introduced for all the images used for training. The first level contains the noise with zero mean and variance in the range of 0.002-0.004, and the second one has a variance of 0.002-0.011 (sample of image with Gaussian noise is shown in Fig. 3).

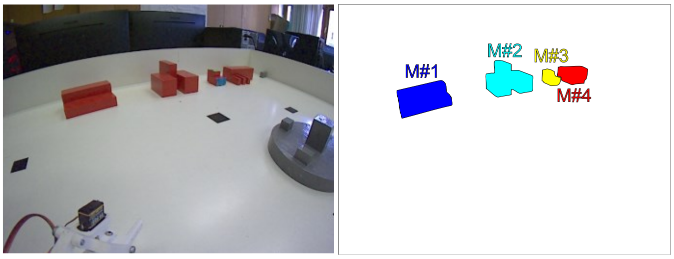


Fig. 2. Sample of the custom-developed dataset

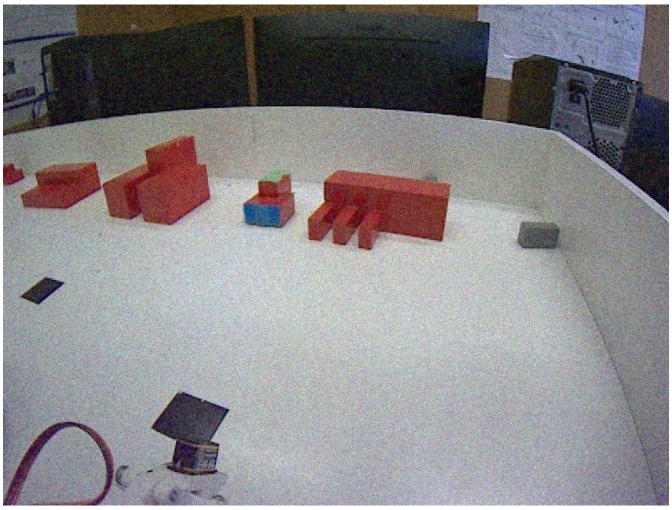


Fig. 3. Image with Gaussian noise

Besides the Gaussian noise procedures, we also applied three data augmentation procedures: (i) horizontal flips (10% of the images), (ii) random crops with a scale of 0.7 (10% of the images), and (iii) the complete image pixel intensities change in the range of 0.8-1.2 (10% of images); this procedure results in images with different illumination intensity levels, which can realistically occur during the different parts of the day. After data augmentation is done, the considered dataset contains 370 images used for training neural networks.

The details about utilized architecture are presented in Fig. 4. Different blocks of layers are presented with rectangles of different colors, while the parameters for those layers are presented within the rectangle. **W** represents the weight matrix dimensions, S is the stride value, and P represents the padding value. Convolution, BatchNormalization, and ReLU layers are presented with blue blocks. The green block presents the MaxPooling layer, while the convolution and BatchNormalization block is presented with the brown rectangle. Finally, the adding layer in combination with the ReLU activation layer is presented with orange. Input images have 800×600×3 resolution, while the output semantic mask has the dimension of 19×25. The probabilities of the class prediction are calculated by utilizing the Softmax activation function (1), while the utilized loss function is Cross-entropy (2).

$$s_i = \frac{e^{y_i}}{\sum_{i=1}^{N} e^{y_i}} \tag{1}$$

$$\ell(\mathbf{s},\mathbf{c}) = -\sum_{i}^{N} c_i \log(s_i) \tag{2}$$

Where **y** represents the output vector of the neural network, i is the current element of the output vector, N is a total number of classes (and the number of elements in the output vector), s_i is the output of the softmax function for each element, **c** represents one-hot vector for the correct class of the current input vector, and ℓ represents the loss function value.

The training is carried out by PyTorch v1.6.0 with Stochastic gradient descend and the momentum of 0.9. The initial learning rate is η=0.01 with the changing schedule defined with (3):

$$\eta^{\text{new}} = \eta^{\text{old}} \cdot \left(1 - \frac{\text{current\_epoch}}{\text{max\_epoch}}\right)^{0.9}. \tag{3}$$

It is important to note that current_epoch is enumerated from 0 to (max_epoch – 1). For the experimental research presented in this paper, the maximum number of epochs is 30, while the mini-batch size is 4. Lastly, the regularization technique is utilized with a weight decay of 0.0001. Training is performed on Nvidia RTX 1660 GPU with 6GB of RAM.

Since the Nvidia Jetson nano is an edge device with limited processing power (NVIDIA 128-core Maxwell GPU), the whole FCN network with encoder and decoder parts could not be implemented in real-time. Therefore, the authors propose to maintain the output of the backbone network and directly calculate the semantic mask with the output resolution instead of deconvolving that information to acquire prediction with the same resolution as input. Having that in mind, the output mask is considerably smaller in resolution than an input image. However, the achieved accuracy is entirely satisfactory.

III. Experimental Results

We have trained the FCN-ResNet18 model on our custom-developed dataset for semantic segmentation of laboratory model of a manufacturing environment. Moreover, the trained model is implemented in the perception system on the mobile robot RAICO. Two metrics utilized to analyze the generalization performance of the FCN-ResNet18 model are Global accuracy and Intersection over Union (IoU). The training results for each class, as well as for the whole dataset, are presented in Table I.

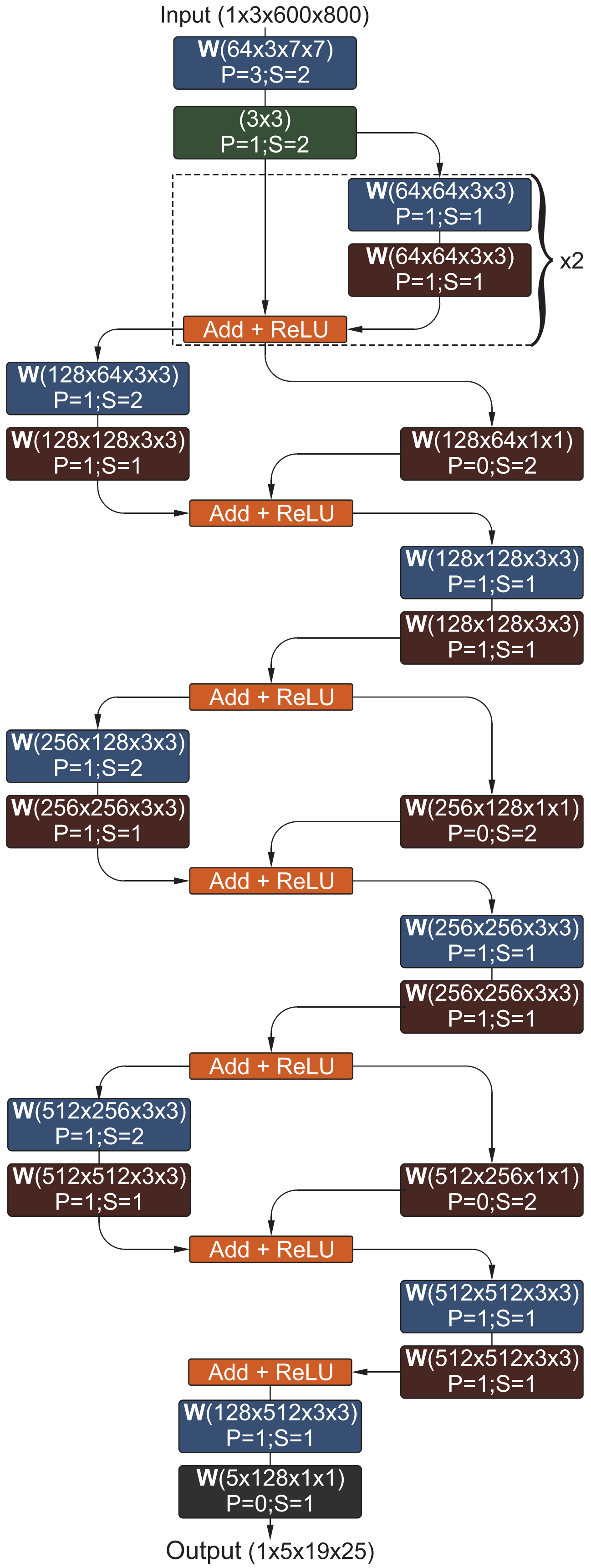


Fig. 4. Architecture of the FCN-ResNet18 model

TABLE I
THE EXPERIMENTAL RESULTS OF THE SEMANTIC SEGMENTATION MODEL

Accuracy measures	Background	M#1	M#2	M#3	M#4
Global per-class accuracy [%]	96.8	68.9	88.6	91.4	94.4
Per-class IoU	96.1	58.2	69.4	45.9	58.6
Mean global accuracy =96.0			Mean IoU = 65.6		

As shown in Fig. 5, there is a significant class imbalance in the considered dataset, as in most semantic segmentation datasets. The dominant class in the images is the background.

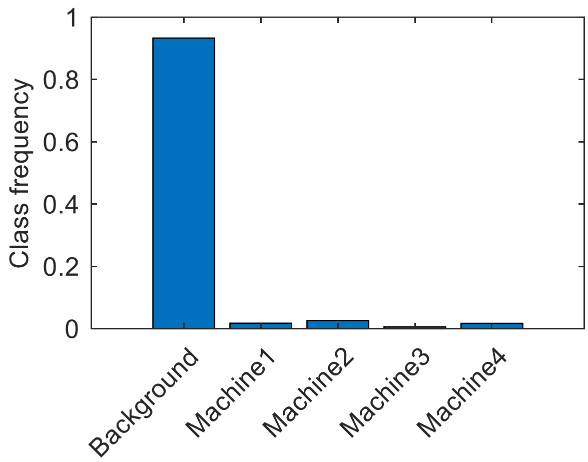


Fig. 5. Class frequency in the custom dataset

Having that in mind, the highest accuracy is achieved for the class with most samples, even though the authors have added the class weights that are inversely proportional to the class frequencies. Moreover, the worst results (for IoU metric) are achieved for Machine3 (M#3) since it is the smallest machine and therefore occupies the smallest percentage of the scene. Interestingly, global accuracy for M#1 is the smallest compared to all the other classes. The authors further investigated this occurrence and presented the overlay view of two test images and their semantic masks generated by the FCN network (Fig. 6). As it can be seen, the network misclassified half of the M#1 in the first image in Fig. 6. Furthermore, in the image with occlusions, part of the M#1 is misclassified and labeled as M#4.

Achieved mean global accuracy is 96.0%, which is a promising result; however, the mean IoU measure of 65.6 is much more representative of the actual generalization capabilities of the FCN model.

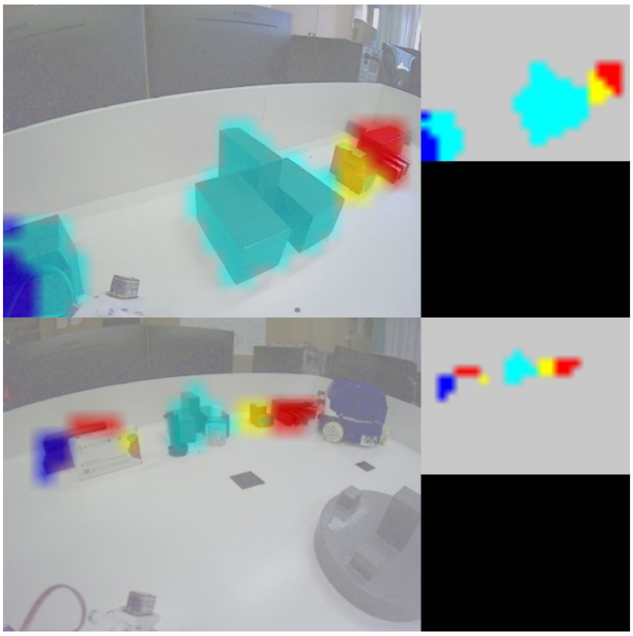


Fig. 6. Test images overlayed with semantic maps

To further test the accuracy of the trained network, the model is implemented on mobile robot RAICO and tested online by the real-time acquisition of images and their semantic segmentation. Fig. 7 presents few images acquired and segmented by the FCN model in a real-world scenario. To increase the effectiveness of the FCN model, it is transformed to an ONNX format and optimized by utilizing Nvidia TensorRT.

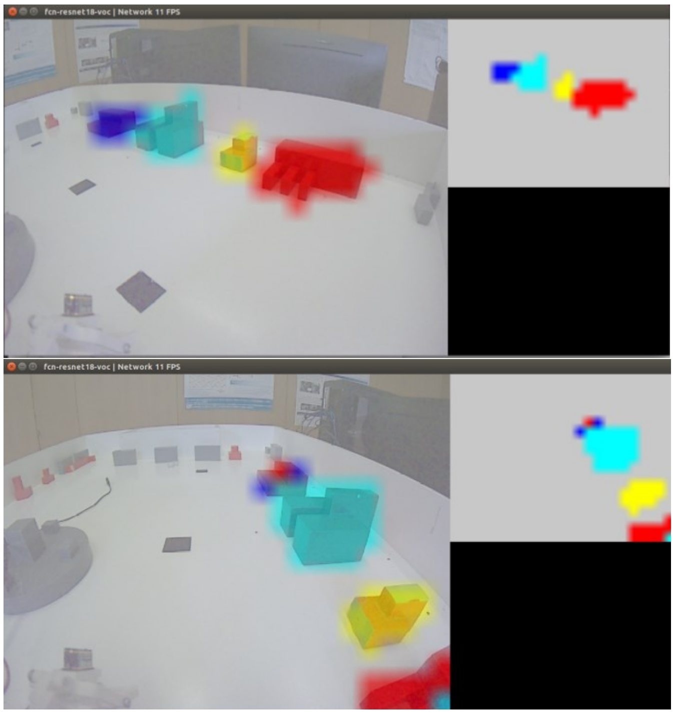


Fig. 7. Testing of implemented FCN model

From Fig. 7, it can be seen that the considered FCN model achieves acceptable accurate segmentation results, with minor errors on machines that are either far away, occluded by other machines, or only partially visible. Furthermore, the model is implemented with 11FPS, which is acceptable for a mobile robot with low-velocity profiles.

IV. Conclusion

This paper proposes the new perception system of mobile robot RAICO based on a Fully Convolutional neural Network with ResNet18 backbone architecture. Training of the neural network model is carried out on a custom-developed dataset for semantic segmentation of the laboratory model of the manufacturing environment. The perception system is integrated with the Nvidia Jetson Nano development board and two Basler dart cameras and configured as a standalone-edge device. After the training procedure is completed, the model is implemented on the mobile robot RAICO, with the achieved accuracy measures of 65.6 for mean IoU and 96.0 for the global accuracy. The implemented system works in a near real-time manner achieving approximately 11FPS. Future research directions could include creating a larger dataset with more classes of manufacturing entities, as well as developing a novel, faster architecture for semantic segmentation capable of running real-time on Jetson nano.

Acknowledgment

This work has been financially supported by the Ministry of Education, Science and Technological Development through the project "Integrated research in macro, micro, and nano mechanical engineering – Deep learning of intelligent manufacturing systems in production engineering" (contract No. 451-03-9/2021-14/200105), and by the Science Fund of the Republic of Serbia, grant No. 6523109, AI – MISSION 4.0, 2020 – 2022.

References

[1] J. Shabbir and T. Anwer, "A survey of deep learning techniques for mobile robot applications," *arXiv Prepr. arXiv1803.07608*, 2018.

[2] T. Ran, L. Yuan, and J. B. Zhang, "Scene perception based visual navigation of mobile robot in indoor environment," *ISA Trans.*, vol. 109, pp. 389–400, 2021.

[3] J. Wu, I. Yildirim, J. J. Lim, B. Freeman, and J. Tenenbaum, "Galileo: Perceiving physical object properties by integrating a physics engine with deep learning," Proc. Advances in neural information processing systems (NIPS 2015), vol. 28, pp. 127–135, 2015.

[4] Z. Yan, S. Schreiberhuber, G. Halmetschlager, T. Duckett, M. Vincze, and N. Bellotto, "Robot perception of static and dynamic objects with an autonomous floor scrubber," *Intell. Serv. Robot.*, vol. 13, pp. 403–417, 2020.

[5] Z. Miljković, N. Vuković, M. Mitić, and B. Babić, "New hybrid vision-based control approach for automated guided vehicles," *Int. J. Adv. Manuf. Technol.*, vol. 66, no. 1–4, pp. 231–249, 2013.

[6] S. Sridhar and A. Eskandarian, "Cooperative perception in autonomous ground vehicles using a mobile-robot testbed," *IET Intell. Transp. Syst.*, vol. 13, no. 10, pp. 1545–1556, 2019.

[7] J. Qian, B. Zi, D. Wang, Y. Ma, and D. Zhang, "The design and development of an omni-directional mobile robot oriented to an intelligent manufacturing system," *Sensors*, vol. 17, no. 9, Article Number: 2073, 2017.

[8] D. M. Chan and L. D. Riek, "Object proposal algorithms in the wild: Are they generalizable to robot perception?," Proc. *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2601–2607, 2019.

Development of applicative interface for connecting optical 3D scanner and robot controller of the UR-5 industrial robot arm

Emilija Stanković, School of Electrical Engineering, University of Belgrade
Ilija Stevanović, Institute Mihajlo Pupin, Robotics department
Aleksandar Rodić, Institute Mihajlo Pupin, Robotics department

***Abstract*— This paper is focused on the implementation, development and testing of an interface for connecting an automated 3D scanner and an industrial robot. Scanning is based on the use of Structured Light scanning software and the accompanying equipment. A rotating platform is constructed, and it is powered by an Arduino-controlled stepper motor. Program in addition to rotating platform by optimal number of degrees, also communicates with the scanning software - via the COM port to synchronize the movement of the rotating platform and the scan. 3D image is obtained by connecting several captured frames during the rotation of the object. The result is point cloud in space, post-processing is performed, and selected points form a robot trajectory. Simulation is performed by MATLAB Robotics Toolbox.**

***Index Terms*— Structured Light scanning, Robotics System Toolbox, UR-5, Point cloud**

I. Introduction

Spatial visualization through realistic 3D modeling is the process of obtaining a mathematical representation of the three-dimensional surface of an object. 3D scanning is widely used in all spheres of industry because it provides fast data collection, and it guarantees a high quality. 3D scanners have the ability to capture in detail points, lines and surfaces of objects, as well as texture and size.

3D scanning is transforming nearly every industry, construction, manufacturing, marketing, medical industry, forensics, reverse engineering in areas of automotive, aerospace and shipbuilding. It has a multitude of applications in robotic vision and guidance, some of them being AI applications, art, design, interior visualization, media, machine control, site layout, flight and wind tunnel testing, prototyping and 3D printing, quality control, and scanning of power equipment assembling of a substation.[1]

Emilija Stanković is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: emilijazstankovic@gmail.com)

Ilija Stevanović is with the Institute Mihajlo Pupin, University of Belgrade, 15 Volgina, 11020 Belgrade, Serbia (e-mail: ilija.stevanovic@pupin.rs)

Aleksandar Rodić is with the Institute Mihajlo Pupin, University of Belgrade, 15 Volgina, 11020 Belgrade, Serbia (e-mail: aleksandar.rodic@pupin.rs)

With that in mind, incorporating 3D scanning technology and robotic manipulators has the potential to facilitate all spheres of industry, from simple things like manufacturing to complicated operations like robotic surgery. Different approaches to using 3D scanners as robotic vision systems have been proposed in the literature.[2][3][4] In industry, robotic systems are taught by guidance to perform specific tasks on object. When changing the object that the manipulator needs to interact with it is necessary to repeat the process of teaching guidance every time. With the introduction of 3D scanning this problem has been avoided and flexibility precision and repeatability are increased. Nowadays, robotic surgery occupies an important place when it comes to the advancement of technology and is being applied more and more every day. Some of the fields are Microsurgery (Micro-manipulation), Skin harvesting (Surface tracking), Neurosurgery Interventional radiology (Constrained targeting). [5]

II. Development goals

Collaborative robotic arms are used in almost all spheres of industry because in addition to classic industrial precision, they also provide safety and protection, so that people can move freely in a collaborative environment.

Direction in which this project is developing, in its primary focus has demonstration that will with the help of new 3D scanning technologies, enable robots to perform functions with great precision. This represents the something that man cannot perform with his free hand. An example of such functions is robotic surgery.

III. 3D scanner hardware description

The system is primarily composed of Logitech C615 web camera (maximum resolution 1920 x 1080 pixels, and range from 30fps to 5fps), Projector - Acer K132 (resolution 1280 x 800 and 500 ANSI Lumens), Structured Light scanning software and calibration pattern and calibration panel. The angle between the walls of the calibration panel is 90 degrees and the size of the calibration pattern depends on the size of the object being scanned.

Additional hardware components required for construction of automatic 360° scanning system for rotating platform, Kinco 2S42Q-0240 Step motor, Kinco 2M412 stepper motor driver, DC Power supply, Arduino UNO, HDMI and USB cables. Algorithm code in Arduino software allows communication between Structured Light scanning software and rotating platform. The optimal angle of the camera in relation to the projector is 22-25º. For objects larger than 350mm the camera is placed to the right of the projector.

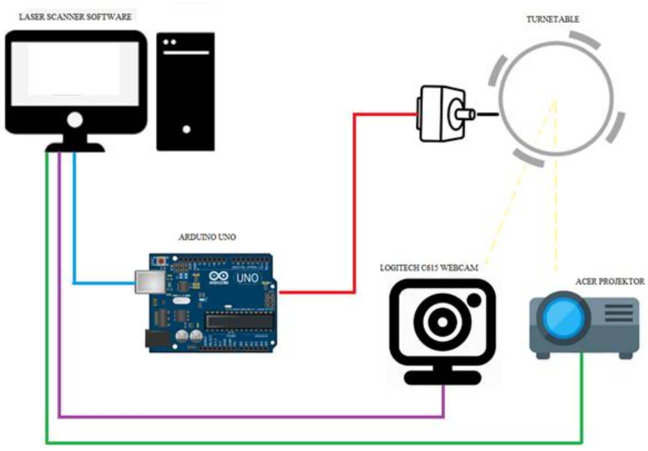


Fig. 1. Hardware configuration of the automatic 3D scanner

IV. DESCRIPTION OF THE APPLICATION INTERFACE

The calibration process consists of two key parts, camera calibration and projector calibration. In order to obtain a projection matrix from the world coordinates to the image in the process of calibrating the camera, the calibration pattern must have at least six points that are clearly visible.[6]

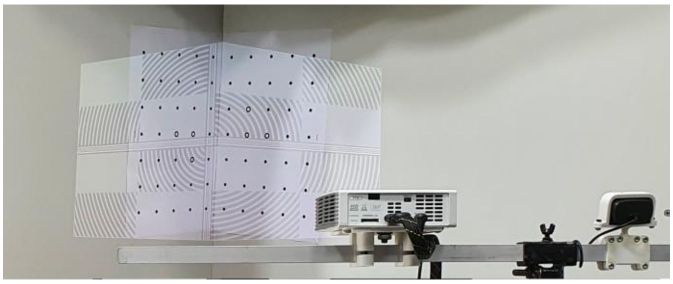


Fig. 2. Calibration process

The projector illuminates a striped pattern, so that software can determine the relative position of the camera and projector. The calibration template has many points, so that in the calibration process the projector can estimate all the unknown coefficients, unambiguously.

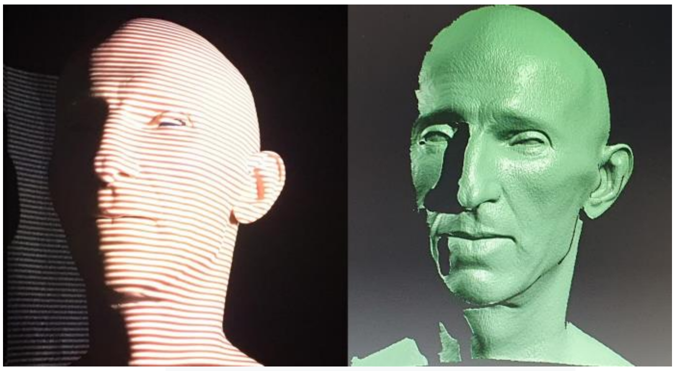


Fig. 3. Scan process (left) and result of one scan (right)

The software has the ability to communicate via the com port by sending numbers and letters. In advanced settings, it is possible to personalize the message that is sent and received.

The commands that the software receives are (Commands):

- Q – AddToList
- I – StartScan
- U – ModeSL
- P – StopLaserScan

The characters that the software sends are (Messages):

- 0 – StartSLScan
- H – GrabImage
- C – ModeSL

An algorithm implemented in Arduino that allows synchronizing movement of the rotating platform and the scanning software:

Serial.available() function returns number of bytes(characters) available for reading from serial port. Input is placed in variable c and the program waits for c to have a value of 0 which represents that the scanning has started. After that, the next character that arrives from the serial port is entered in the variable p, as long as p does not get the value of H, which means that the frame is captured, it is reloaded. When p is equal to H, the algorithm waits for msAfterPictureTaken. The value of this time content depends on the complexity of the scan, the image quality, and whether it is necessary to collect the texture color. A Q character is then sent to the serial port, which is a command to the software to save the current scan in a List. Program then waits for msAfterSave, and the platform on which the object is located rotates by Rotation angle. The Rotation angle value is determined before the scan begins and depends on the complexity of the object. In order for the platform to stabilize after the change of position, it is necessary to wait for msAfterRotation, the cnt counter is incremented by one and the U command is sent to the software to prepare for further scanning, to adjust all camera and projector parameters. If the value of the cnt counter is less than the number n obtained when 360 is divided by the Rotation angle, full rotation is not completed and scanning is continued by sending an I character to the serial port indicating the command for the software to start scanning. If a full rotation is performed, the command P is sent the scanning process is stopped and thus the process of collecting individual scans is completed.

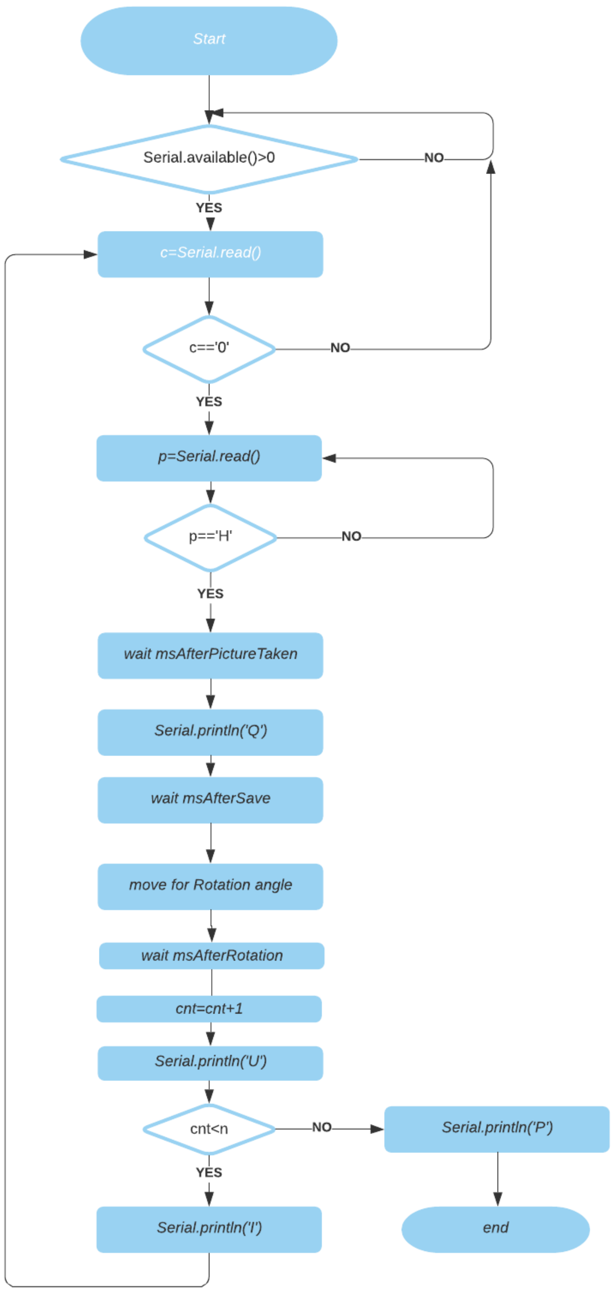


Fig. 4. Flowchart of the algorithm for synchronizing movement of the rotating platform and the scanning software

The duration of the scanning process depends on the quality of the scanned object required for proper application, so it needs to be determined experimentally. After collecting N scans, they are saved and exported to MeshLab to be aligned and then fused.

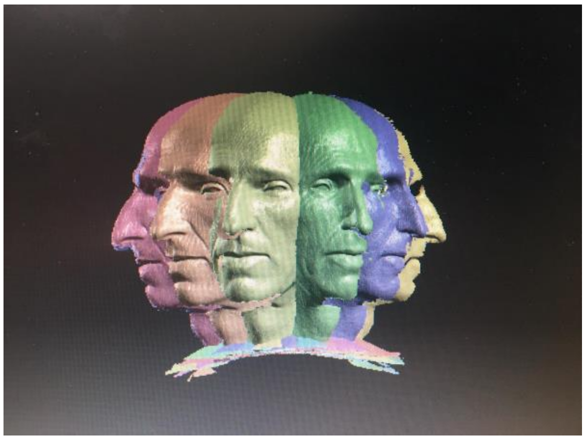


Fig.5. Screenshots of the succesive scans

The result of the scan is the obj file format – it represents 3D geometry of the object. File contains:

- v-vertices 3xN matrix contains x, y, z coordinates of points.
- vt- texture vertices optional if texture capture is enabled during scanning.
- vn- vertex normal.
- f- faces (face is any of the individual flat surfaces of a solid object)

For files of this complexity to be used in the simulation in MATLAB, it is necessary to simplify the scan result and for that the MeshLab software package was used.

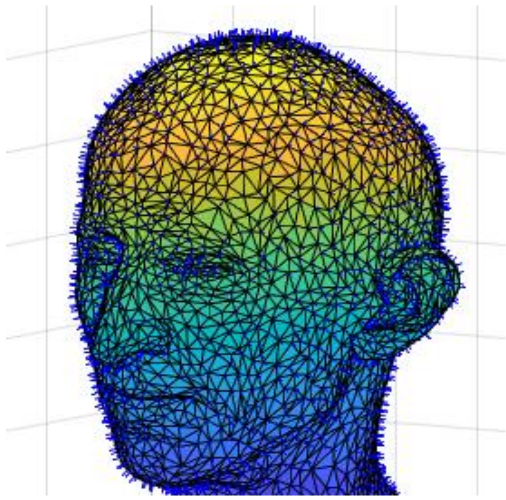


Fig.6. Mesh representation – of simplified scan in MATLAB

The position vectors consist of the x, y and z coordinates of each vertex. Let n_α be the normal vector derived from the mesh file, the components of the rotation matrix of the end effector are obtained as follows [8]:

$$\hat{a} = -\frac{n_\alpha}{\|n_\alpha\|}, \hat{n} = \hat{a} \times \begin{bmatrix}0\\1\\0\end{bmatrix}, \hat{o} = \hat{a} \times \hat{n} \quad (1)$$

$$R_{EE} = \begin{bmatrix}\hat{n} & \hat{o} & \hat{a}\end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33}\end{bmatrix} \quad (2)$$

From there we can derive values that represent the orientation of each vertex:

$$\begin{aligned}\psi &= atan2(r_{21}, r_{11})\\ \theta &= atan2(-r_{31}, cos(\psi)r_{11} + sin(\psi)r_{21})\\ \varphi &= atan2(sin(\psi)r_{13} - cos(\psi)r_{23}, -sin(\psi)r_{12} + cos(\psi)r_{22})\end{aligned} \quad (3)$$

V. CONNECTION OF SCANNER AND ROBOT CONTROLLER

Universal robots have the ability to communicate over TCP/IP protocols (TCP socket connection over Ethernet). A PC is a server while a robot is a client. It is possible to establish communication by writing scripts in different programming languages, but as the simulation was performed in MATLAB the communication is also implemented in MATLAB. MATLAB has built-in functions for creating servers and clients. It is necessary to know the IP address and port number. Server waits for client connection. The robot needs to send the message that the server expects in order to get the first desired waypoint as feedback.

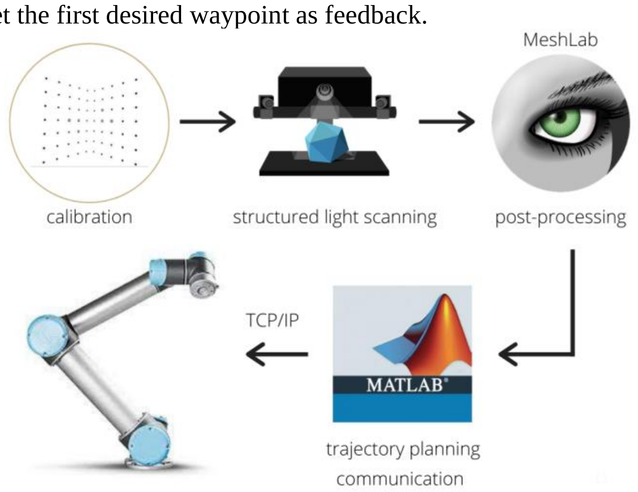


Fig.7. High-level description of the integrated system - 3D scanner and industrial robot UR-5

The message the robot receives consists of the number of data sent and that data, so that at any time it is possible to check whether the message has been sent in its entirety (format in PolyScope: *receive_data≔[6,0,0,0,0,0,0]*).

Data sent from server must be in bracket, each data must be followed by a comma unless it is the last, at the end there is '/n'. The robot has three main ways of calculating how to move from Waypoint to Waypoint which is a nonlinear movement “MoveJ”, a linear movement, “MoveL” and a circular movement (MoveC) which is under a Process move “MoveP”. [9]

The Universal Robot is controlled on Script Level, URScript is the robot programming language (PolyScope software).
An example of the functions used [10]:

- *Movel(pose, a=acceleration, v=speed,r=blend radius)*
  move to position (linear in tool-space), a[m/s2], v[m/s], r[m].
- *get_forward_kin()*
  returns tool pose-forward kinematic transformation (joint space to tool space) of current joint positions.
- *get_inverse_kin(x)*
  returns joint position- Inverse kinematic transformation (tool space to joint space). Solution closest to current joint positions is returned.
- *socket_open(server, port)*
  Open ethernet communication.
- *socket_read_ascii float(number)*
  Reads a number of ascii float from the TCP/IP connected.
- *socket_send_string(str)*
  Sends a string to the server Sends the string through the socket in ASCII coding

Robot sends current tool positions (fun. *socket_send_string*) in the format: p[x,y,z,ψ,θ,φ]. Functions for reading current positions of the final device are written in MATLAB.

VI. DEMONSTRATION EXAMPLE

Experimental verification of the system was performed by scanning the model of Nikola Tesla's head.

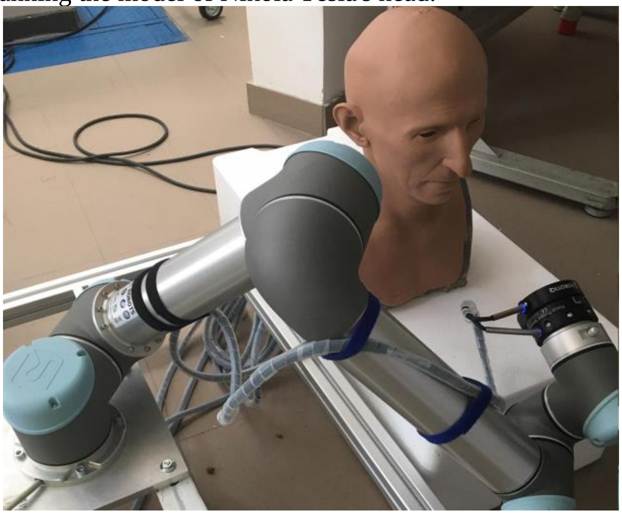


Fig8. The interaction of the robotic arm and the model of Nikola Tesla's head

In order to better simulate the movement of the robotic arm, the points obtained as a cross section of the point clouds and horizontal plane passing through the axis of symmetry were chosen as a demonstration example.

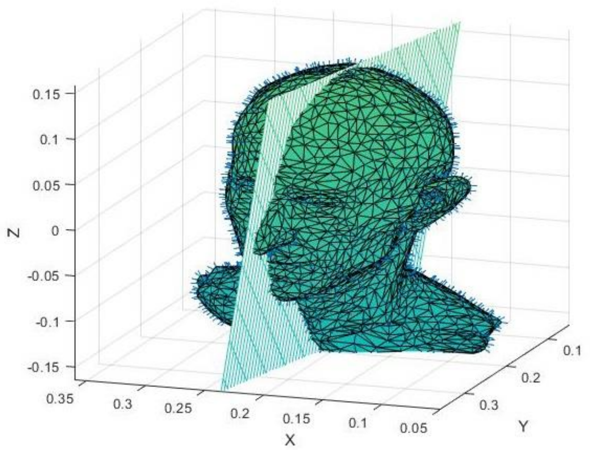


Fig.9. Intersection of planes and mesh representations

The distance of each point from the plane was calculated and based on that, points with a smaller distance from the threshold were selected. A value of 5mm was chosen for the threshold and *Figure10.* shows which points meet this criterion.

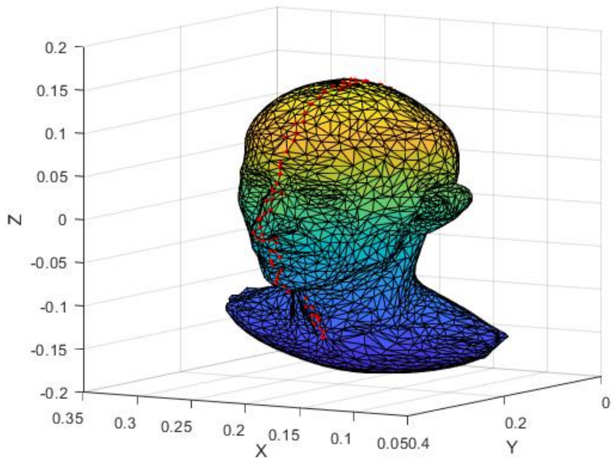


Fig10. Selected points

The robot from the initial position crosses the contour of the face at a constant speed of one cm per second, continuously without stopping.

A laser is placed on the end effector for easier visualization. In order to find a point on the line(X) containing the normal vector(n) that passes through the original point(P) on the scanned head model and is d=1 cm away from the original point, the following system of equations is solved:

$$\vec{n} = \begin{bmatrix} l \\ m \\ s \end{bmatrix} \qquad \begin{aligned} P &= (x_1, y_1, z_1) \\ X &= (x, y, z) \end{aligned} \tag{4}$$

$$t = \frac{x - x_1}{l} = \frac{y - y_1}{m} = \frac{z - z_1}{s} \tag{5}$$

$$d = \sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2} \tag{6}$$

The result of the displacement of the profile points by 1 cm in the direction of the normal is shown in the *Figure11*.

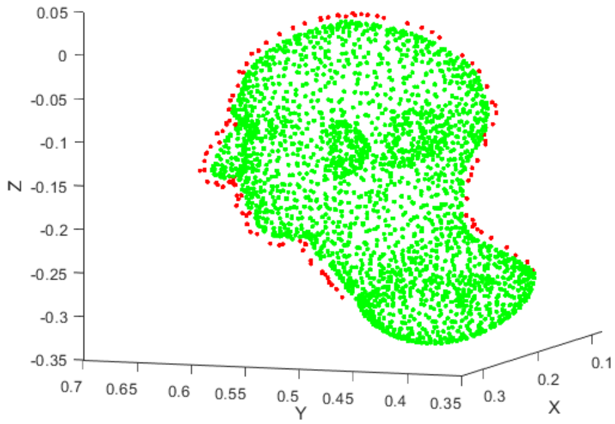


Fig.11. New selected points

For the purposes of simulation, the UR5 robot was imported from the Robotics System Toolbox, as rigid Body Tree.

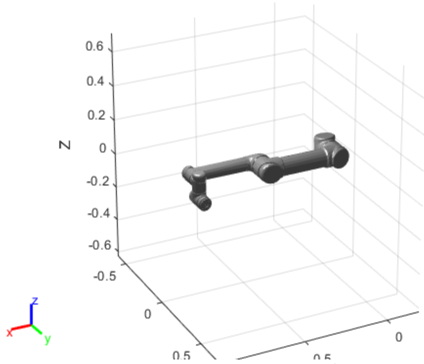


Fig.12. Robot UR5 in home configuration position

By calling the *inverseKinematics* function, an inverse kinematics object is formed which, based on the desired position and orientation of an end-effector computes the joint configuration that leads to it.[11] The numerical optimization used in the calculation of inverse kinematics is the Broyden – Fletcher – Goldfarb – Shanno algorithm (BFGS):

Taking into account the starting point(q_0) and the initial Hessian matrix(H_0) for each iteration(k), the search direction(d_k) is calculated:

$$d_k = -H_k g_k \tag{7}$$

If the gradient(g_k) is zero, the search stops. If not, step size (α_k) is calculated with line search method, and new point is obtained:

$$x_{k+1} = x_k + \alpha_k d_k \tag{8}$$

Then the new Hessian matrix is recalculated, and the stop criterion is examined. If it is not fulfilled, it moves to the next iteration.[12]

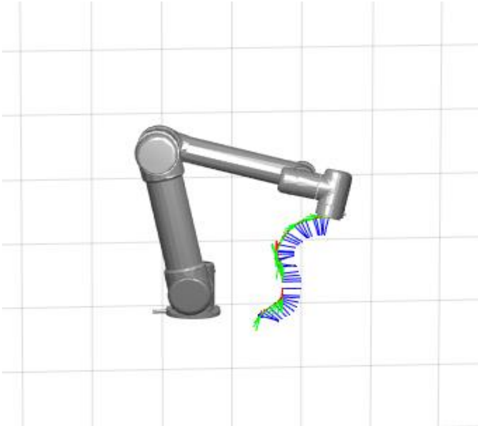


Fig.13. Simulation of trajectory monitoring

Motion trajectory was implemented as Joint space and Task space trajectory. In addition to interpolating the positions between each waypoint, it is necessary to implement interpolation between orientations. This is achieved by interpolation between quaternions owning to the fact that if we interpolated the Euler angles the solution would not be unique. The way this is implemented in MATLAB is the SLERP (Spherical Linear Interpolation) method.

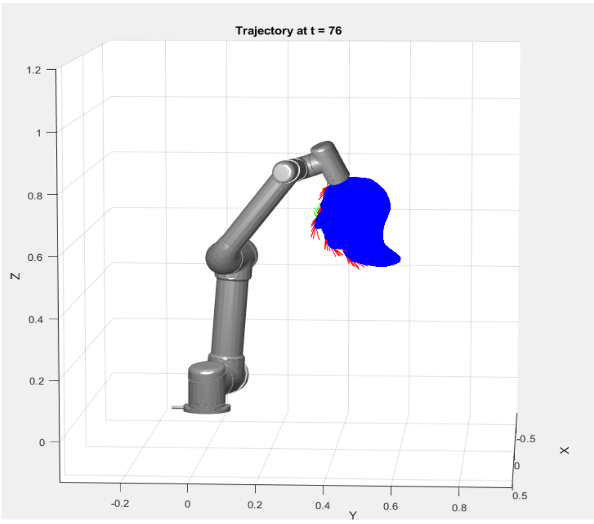


Fig.14. Simulation of trajectory monitoring

Another way to obtain the internal angles of the robot from the external coordinates of position and orientation is to implement an analytical procedure for solving inverted kinematics.[13] As there are several configurations that can produce a certain position and orientation, one of the solutions is to choose the configuration that makes the smallest change in the position of the joints compared to the previous point. This method is much slower, so numerical calculation is applied with the help of libraries and built-in functions in the MATLAB Robotics Toolbox.

VII. Conclusion

Automatic 3D object scanning facilitates processes not only in robotics but also in other important spheres of life. Collaborative robots and their integration with such systems provide a new form of technological advancement that aims to improve the way people live and work. The results of the scanning of Nikola Tesla's head met the criteria of precision. The goal of this work, the construction of automated scanning and connection with the universal robot UR5 is achieved. Automated scanning is implemented using simple hardware for the motorized movement, and low cost or open-source software. Moreover, the process of filtering and collecting point clouds that form the robot trajectory is described. Experimental verification of the system was performed and communication is established.

Acknowledgment

The results presented in the paper are obtained in the scope of the research projects: „Development and Experimental Performance Verification of Mobile Dual-Arms Robot for Collaborative Work with Humans “, Science and Development Programme – Joint Funding of R&D Projects of the Republic of Serbia and the People's Republic of China, contract no. 401-00-00589/2018-09, 2018-2021 and national R&D project no. TR-35003, both supported by the Ministry of education, science and technology development of Republic Serbia.

References

[1] O. Vozisova, S. Eroshenko, E.Koksharova, A. Khalyasmaa and S. Dmitrir (2016, March), “Application of 3D Scanning and Printing Technologies in Electric Power Industry,” Ekaterinburg, Russia. In 2016 IEEE International Conference on Industrial Technology (ICIT) (pp. 892-897). IEEE

[2] S. Koceski, N. Koceska, P. B. Zobel and F. Durante, "Characterization and modeling of a 3D scanner for mobile robot navigation," 2009 17th Mediterranean Conference on Control and Automation, 2009

[3] E. Digor, A. Birk, A. Nüchter “Exploration Strategies for a Robot with a Continously Rotating 3D Scanner”, Simulation, Modeling, and Programming for Autonomous Robots, 2010, Volume 6472

[4] Ogorodnikova, O., Olchanski, D. “On a 3D scanning robot system design problem”, Periodica Polytechnica Mechanical Engineering, 51(1), pp. 39–44, https://doi.org/10.3311/pp.me.2007-1.06

[5] Etienne Dombre „Introduction to Surgical Robotics“, LIRMM, Montpellier 2009.

[6] Josep Forest i Collado, “New methods for triangulation-based shape acquisition using laser scanners,” *Girona, maig 2004*.

[7] Du Q. Huynh, “Calibration of a Structured Light System: A Projective Approach,” Centre for Sensor Signal and Information Processing Signal Processing Research Institute Technology Park, The Levels, SA 5095.

[8] Jelena Z. Vidaković, Vladimir M. Kvrgić, Mihailo P. Lazarević, Zoran Z. Dimić, Stefan M. Mitrović (2017), “Procedure for Definition of End-effector Orientation in Planar Surfaces Robot Applications,” Belgrade Tehnika, 72(6), 845-851.

[9] UNIVERSAL ROBOTS User Manual, Version 3.3.3.

[10] The URScript Programming Language Manual.

[11] Peter Corke, Robotics Toolbox for MATLAB, Release 9.10, February 2015. https://petercorke.com/toolboxes/robotics-toolbox/

[12] Atikah Ramli, Ibrahim Jusoh, Mohd Rivaie Mohd Ali (2017), “A combination of Broyden-Fletcher-Goldfarb-Shanno (BFGS) and n-th section method for solving small-scale unconstrained optimization”, Malaysian Journal of Fundamental and Applied Sciences, 13(4), 717-720.

[13] Ryan Keating, “UR5 Inverse Kinematics”,Johns Hopkins University.

Fusion of Camera-Acquired Data and CAD 3D Models of Objects in Forming a Visual Feedback Loop for Industrial Robots

Miloš Nenadović, Uroš Ilić, and Aleksandar Rodić

Abstract—**In this paper, we present a computer vision system for object detection and spatial position and orientation recognition. To solve the problem, we separated the process in several stages: in the first stage the system uses principal component analysis (PCA) method in optimal conditions to detect objects in the image. Then, after the object is extracted, we projected it in the appropriate eigenspace, produced by singular value decomposition (SVD) of the set of images of the rotated 3D CAD model. The closest match is then processed by correlation between it and the real-object image in log-polar space. The result is combined with information from other cameras to derive the approximate position and object orientation using multiple-view geometry.**

Key words—**Object detection; Singular Value Decomposition; Log-Polar; Multiple-view geometry.**

I. Introduction

The problem of object detection in images is one of the oldest in computer vision [1] and has over the years been solved by various methods, among which are neural networks and principal component analysis (PCA). Convenient in the sense that they only require a set of training data to extract key features from objects, they provide vastly accurate classification results. However, as the accuracy rises so does the required memory, which is also the case as the complexity of objects rises.

In industrial robotics we require the robot arm to handle objects of varying complexities, ranging from simple boxes to machine parts. In these cases, a camera is a good tool to use to scan the scene and retrieve information about objects. The problem arises when we need to do more than classify an object. Assuming we are only provided with several cameras, we aim to get the most information about the scene which would be the position and orientation of objects.

We will aim to decrease the amount of images necessary to train the neural network, or in the case of principal component analysis [2] to eliminate the need for them altogether and instead only use the 3D CAD models provided. The CAD models are used to obtain reference images of their respective objects in different positions [3]. From these images we calculate the histogram of oriented gradients (HOG) and use them for singular value decomposition (SVD) based on the idea that these features are enough to determine, at least the approximate, orientation of an object. Following this, we will use correlation of gradient data in log-polar space which is rotation and scale invariant. Additionally, we will analyze the speed of the algorithm to determine whether it can be used for moving objects. We will test the system on a set of simple objects (cube, cuboid, cylinder and pyramid) of different colors (red, blue, green, brown, orange, black). In future work, we will attempt to use the system on more complex objects such as machine parts or 3D wooden puzzle pieces.

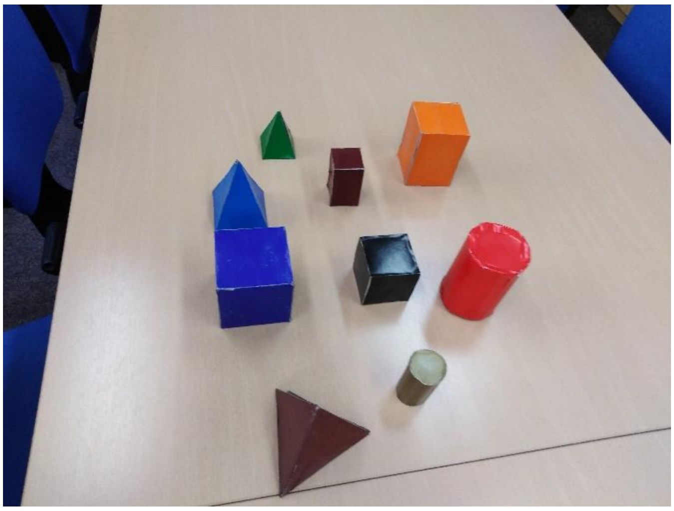


Fig. 1. Scene setup with simple objects of multiple colors on a table.

II. The Role of Robot Vision and 3D CAD Modeling in Trajectory Planning

To understand the need for proper object position and orientation recognition, we look at an example of a robot grappler arm. We assume the objects are still and placed on a table in various positions. To successfully grab an object, the robot arm must be placed so that the object does not slip from its grasp. Additionally, a robot may need to navigate between objects and in such cases the knowledge of their position and

Miloš Nenadović is with the Faculty of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: neosmilos@gmail.com).

Uroš Ilić is with the Faculty of Mechanical Engineering, University of Belgrade, 16 Kraljice Marije, 11120 Belgrade, Serbia (e-mail: ilicuros01@gmail.com).

Aleksandar Rodić is with the Institute Mihajlo Pupin, 7 Sazonova, 11000 Belgrade, Serbia (e-mail: aleksandar.rodic@pupin.rs).

orientation is crucial [1]. Camera images can be used to determine the approximate position of an object using multiple-view geometry, after knowing the exact or approximate positions of the cameras. In this case we assume camera positions are unknown except for their general relative position i.e. 'front', 'left' and 'above'. Thus, we used reference points to determine the approximate camera positions before using them to determine the object positions.

III. Configuring the 3D Robot Vision System

Although some results are possible without calculating the positions of the cameras, we strive to achieve the most accurate representation of the scene and so require this as the first step. Given that no prior information is known about the position and angle at which the cameras observe the scene, we use a set of reference points on the scene to determine them. The size, position and distance between the reference points is known. We search the images for these points, assign the appropriate labels for them, and then, using the known data we reconstruct the camera position and orientation.

IV. Fusion of 3D CAD Model Data and Camera Images

The CAD models are used to construct a set of images from various views by applying rotation transformations to them. For the SVD algorithm we use rotations around each of the axes in increments of 15 degrees, whereas for the log-polar correlation we use 5-degree increments, however for only two axes. This is because one of the axes contributes solely to 2D image rotation.

As for the vision data, as mentioned before, we use a set of three cameras to observe the scene. Each camera image is processed on its own in parallel with the others and upon completion updates the other two. In this way, should any image be wrongly paired, the image with the higher correlation coefficient will take precedence and change the initially estimated view-image to another one based on the camera positions.

In this way we achieve a feedback loop by utilizing CAD model data to interpret the information acquired by cameras and then comparing them to the real thing.

V. Recognition of Object Shape and Spatial Orientation Using Images and 3D CAD Models

A. *Object detection using PCA*

Before the object orientation algorithm can begin, we must first detect objects in the image. One of the goals we set ourselves was to detect the objects without using real-object images, but rather just the models. For this purpose, we created 936 training images for each of the models, depicting different views that we used as the training set.

The PCA method for object detection relies on using a training set to find principal components (PCs) to use to reconstruct images. Assuming that the best image reconstruction can be done only when using the PCs created from an image set of the same type [2], based on the accuracy of the reconstruction, we can classify an image as one of the objects or part of the background.

As mentioned in [2] we use vectorized reference intensity images to acquire a projection matrix P, and μ as the mean of the set. When given a vectorized intensity input image u we get the reconstructed image r through:

$$r = P'P(u - \mu) + \mu. \quad (1)$$

The reconstructed image is then compared to the original image and their difference d is expressed as the 2-norm between the two image vectors, as in

$$d = \|r - u\|. \quad (2)$$

To test this method, we supplied an image of the blue pyramid object as the input image and then used different training sets containing reference model-images and looked at differences d they produced.

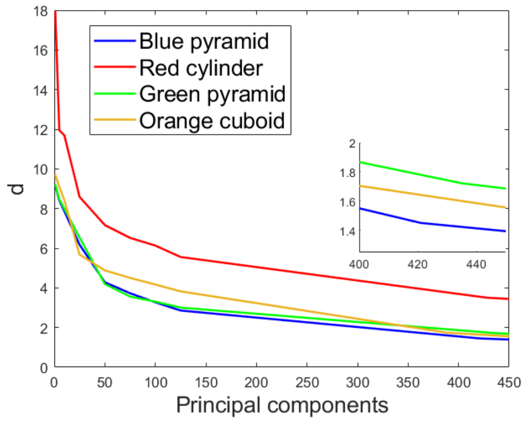


Fig. 2. Image reconstruction of the blue pyramid image using principal components obtained from four different model data sets (blue pyramid, red cylinder, green pyramid, and orange cuboid).

As shown in Fig. 2, the best reconstruction is achieved by using the appropriate training data set. We used a set of fixed numbers for the number of PCs used to reconstruct the image, except for the final one, which was the number of PCs that conserved 90% of the singular energy α, similar to how it was done in [3].

One of the problems this method encountered was the situation when, instead of an object, the input image was that of the background. Given its simplicity in comparison to an image containing an object the difference d was even lower than those of the desired object. Thus, we found that it was necessary to have a minimum number of real-object images and background images to create an additional classifier, or for the parts of the image with just the objects to be extracted before applying PCA.

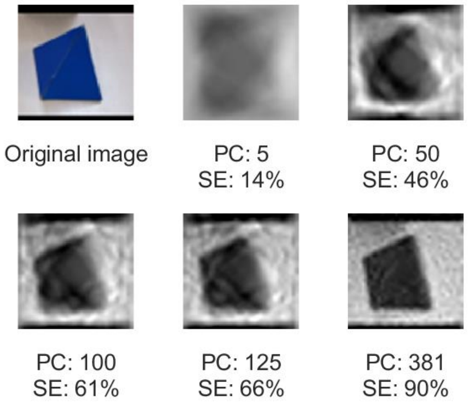


Fig. 3. Image reconstruction using principal components obtained from the blue pyramid model data set, with PC showing the number of principal components used and SE representing the retained percentage of singular energy.

As shown in Fig. 3 increasing the number of PCs also increases the accuracy of the reconstruction. However, doing so beyond a certain point would reduce efficiency instead as the increase of accuracy does not justify the decrease in computation speed and added memory.

B. Finding the closest match using the SVD algorithm

After extracting the object from the entire image, we proceed to use singular value decomposition to find the closest match among a set of rotated model images. As an object's orientation can be represented by its edges, we use gradient analysis to find features. The histogram of oriented gradients provides us with the necessary data t (column matrix of size $n \times 1$) that we place as the columns of a matrix $T = [t_1\ t_2\ \dots\ t_N]$, which the SVD algorithm is applied to.

$$T = U\Sigma V^T \tag{3}$$

From (3) we take the matrix U and use the first k ($1 \leq k \leq N$) columns representing eigenvectors of the largest k eigenvalues in S, sorted in descending order [2]. Like so, we derive the matrix U' from U. Multiplying with vectorized HOG features of input image data u we can project that data to the eigenspace, while reducing dimensionality. This is also applied to all columns of the matrix T creating a matrix T' of reduced size $k \times N$. We then compare the projected data of the input image u with the other projections t' from reference image HOG features (columns of T'), as in

$$d_i = \left\| t_i - U^{T'} u \right\|. \tag{4}$$

Index i in (4) denotes the index of the projection being compared with the projected input image. We store the results in an array which we sort at the end of the algorithm, and the index with the minimum difference points to the reference model image that would be the best pair to the input image.

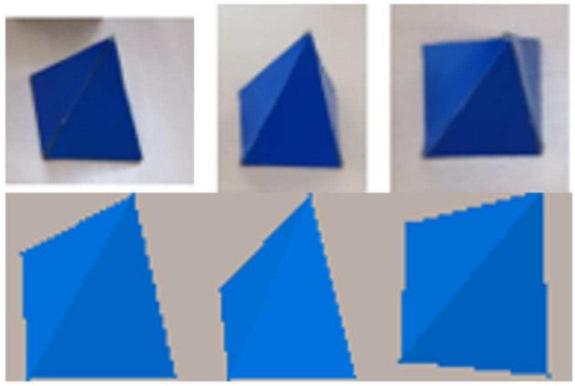


Fig. 4. Result of SVD comparison between extracted object images (top row) from three views (front, left and above) and a set of rotated model images (bottom row).

We can see in Fig. 4 that the algorithm provides two good matches and one bad match. Additionally, we can see that the sides of the pyramid in bad lighting were omitted from the approximations. However, as the algorithm's main purpose is just to narrow the search range for the subsequent parts, having even two out of three bad matches wouldn't pose a problem.

In future work, we will focus on improving the quality of the SVD method to include more accurate estimates both in optimal conditions, as was done in this paper, and when there is occlusion as the total time required for the match to be found greatly depends on the initial estimate.

C. Finding the closest match using log-polar correlation

Aside from the SVD algorithm used to narrow down the possible choices for object positions, we use intensity image correlation in log-polar space to deduce the best pairing of the real-object-image and view-image of the model. Log-polar space is used as it is both rotation and scale invariant [5]. Correlation is used to find the pair with the highest correlation coefficient and determine the rotation angle which is proportional to the vertical shift.

Fig. 5. Model images from multiple viewpoints (first row) and the log-polar transformations of the paired intensity images (second row)

One thing that can be seen from Fig. 5 is that all the log-polar transformations are quite unique. As such, it is not required to store all the information from those images. Rather, as they are computed from the center of the object and images encompasses mainly the object in question, we can take an area of interest that is the second half of the log-polar image along the ρ-axis. This area contains the entire object outline as well as the endpoints of the object edges within the outline.

Problems that may occur in this stage are related to the center used for computing the log-polar transformation. Because we were dealing with objects without occlusion and that are distinct from the background, we could use simple methods of color segmentation to find all the object pixels and then determine the center as a "center of mass".

However, in cases where occlusion is present, or the object is more complex, or the object detection algorithm does not perfectly capture the object, finding the exact center becomes a problem. As mentioned before, log-polar correlation is both rotation and scale invariant, but it is sensitive when it comes to the center from where it is calculated. In case the estimated center of the object is shifted from where it is located on a template image then correlation may not yield adequate results, depending on how much the center is shifted. In [5] it is shown that the best correlation coefficient is achieved when the location of the object center matches that of its template image.

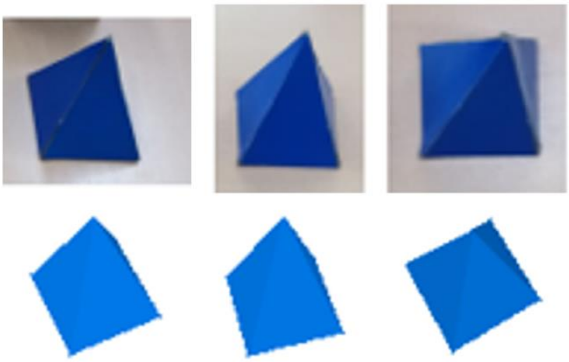


Fig. 6. Real-object images (first row) and best pairs from log-polar correlation without angle correction (second row).

The results shown in Fig. 6 show that the log-polar correlation method finds accurate matches for the objects in question. This means that it is safe to use the less-accurate SVD method first to provide an initial object orientation estimate.

VI. Experimental Results and Discussion

The algorithm was run under the assumption of real-time work, meaning that the number of additional angles around the ones found by the SVD method, which would be checked by the log-polar correlation, had to be lowered to two sets of 4 angles on each side of the central one found by SVD (total of 9 angles). With such modifications the algorithm ran at around 60 ms for single-object orientation detection. This is excluding the time for object detection which took around 50 ms. We will attempt to further bring down the required computation time for this task in future work.

Additionally, one of the possible alterations that we tried was lowering the number of angles that would be checked to only one, the one provided by the SVD method. This lowered the computation time from 60 to around 10 ms.

The possibility of larger errors occurring due to SVD was taken into account and thus after every iteration (determining the 6D position of every object in the scene), the results are carried over to the next iteration until a plateau is hit. Based on the experimental results, we have determined that the algorithm is capable of operating in real-time.

The entire algorithm we proposed is an iterative one and can be summarized in several steps:

1. We use PCA to detect objects in the image for each of the cameras. This step can be skipped after the first run in cases where the scene is static.
2. We calculate an input image's HOG features and send them to the SVD algorithm to estimate the object's orientation. This is done for every camera image, resulting in three estimates.
3. The estimates are then turned to intensity images, transformed into log-polar space, and processed by 2D correlation. The resulting coefficients are compared, and the best result is sent further.
4. Knowing the transformation matrices between cameras, we use the estimate to obtain images that the other cameras should see. These images are treated as the new input images and sent to Step 1.

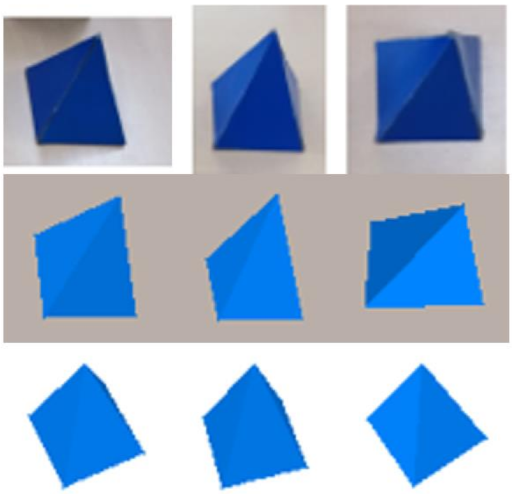


Fig. 7. Iteration 1 of the algorithm. Real-object images (first row), SVD estimated object view images (second row) and log-polar correlation results without angle correction (third row).

Fig. 7 shows that even in the first step of the algorithm we can obtain accurate results, similar to ones in Fig 6 and that the algorithm would end in the second iteration. The total number of iterations needed for convergence varies depending on the initial orientation estimations as well as the number of camera images that we can work with. In cases where one or two of the cameras cannot see an object, that number would exponentially rise.

VII. Conclusion

In this paper, we have presented a method of determining the 6D position of an object based on singular value decomposition and correlation of gradients in log-polar space. PCA allowed us to accurately detect an object, while using the same method with a set of HOG features allowed us to give an initial estimate of the object's orientation. Afterwards, we have taken the gradients of the grayscale images of the estimates and transformed them to log-polar space. The properties of the log-polar space, scale and rotation invariance, have allowed us to lower the number of matches (number of reference images) to compare, reducing computation time. Moreover, we have taken only part of the log-polar images, further reducing computation time. Finally, using iterations assures operation in real-time, while the results on Fig. 6 show that the algorithm will converge with accurate results.

ACKNOWLEDGMENT

The results presented in the paper are obtained in the scope of the research projects: “Development and Experimental Performance Verification of Mobile Dual-Arms Robot for Collaborative Work with Humans”, Science and Development Program – Joint Funding of R&D Projects of the Republic of Serbia and the People’s Republic of China, contract no. 401-00-00589/2018-09, 2018-2021 and national R&D project no. TR-35003, both supported by the Ministry of education, science and technology development of the Republic of Serbia.

REFERENCES

[1] Labbé, Y., Carpentier, J., Aubry, M., & Sivic, J. (2020, August). “CosyPose: Consistent multi-view multi-object 6D pose estimation”. In *European Conference on Computer Vision* (pp. 574-591). Springer, Cham.

[2] Luis Malagón-Borja, Olac Fuentes, “Object detection using image reconstruction with PCA”, *Image and Vision Computing, 27 (1-2), 2-9*

[3] Huang, S. C., Huang, W. L., Lu, Y. C., Tsai, M. H., Lin, I. C., Lau, Y. C., & Liu, H. H. (2019, January). “Efficient Recognition and 6D Pose Tracking of Markerless Objects with RGB-D and Motion Sensors on Mobile Devices”. In *VISIGRAPP (1: GRAPP)* (pp. 375-382).

[4] H. S. Prasantha, H. L. Shashidhara and K. N. Balasubramanya Murthy, "Image Compression Using SVD," International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007), 2007, pp. 143-145

[5] R. Matungka, Y. F. Zheng and R. L. Ewing, "2D invariant object recognition using Log-Polar transform," 2008 7th World Congress on Intelligent Control and Automation, 2008, pp. 223-228

Influence of muscle co-contraction indicators for different task conditions

Marija Radmilović, Djordje Urukalo, Miloš Petrović, Filip Bečanović and Kosta Jovanović

***Abstract*— In this research paper, arm co-contraction indicators are examined in different scenarios such as load variation, hand velocity variation, and in tasks with different precision. The experimental results show the relationship between muscle co-contraction and increase of load, velocity, or precision. According to the results, the differences in the muscle co-contraction related with gender and the age of the participants for the same task are evident. The results of the analysis for each task are in the align with the results presented in the previous research. The results of this research have made a significant contribution in analyzing human stiffness and its implementation in the human-like motion of robots.**

***Index Terms* — muscle, co-contraction, musculoskeletal stiffness, biceps, triceps, anterior deltoid, posterior deltoid**

Marija Radmilović is with the Mihailo Pupin Institute, University of Belgrade, 15 Volgina, 11060, Belgrade, Serbia, (e-mail: marija.radmilovic@pupin.rs).

Djordje Urukalo is with the Mihailo Pupin Institute, University of Belgrade, 15 Volgina, 11060, Belgrade, Serbia, (e-mail: djordje.urukalo@pupin.rs).

Miloš Petrović is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: petrovic.milos@etf.bg.ac.rs)

Filip Bečanović is a joint Ph.D. student with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia, and the University of Paris-Est Créteil, Laboratory of Images, Signals and Intelligent Systems, 120 rue Paul Armangot, 94400 Vitry sur Seine, (email: becaphilippe@gmail.com)

Kosta Jovanović is with the School of Electrical Engineering, University of Beglrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: kostaj@etf.bg.ac.rs)

Distribution of Control Tasks to Smart Devices in Industrial Control Systems: a Case Study

Zivana Jakovljevic, *Member, IEEE* and Dusan Nedeljkovic

***Abstract*—Cyber Physical Systems (CPS) and Internet of Things (IoT) open the way for new generation of Industrial Control Systems (ICS) characterized by high flexibility, modularity and reconfigurability necessary within Industry 4.0. Inevitable shift from centralized to distributed control systems is underway, but the changes are not as rapid as expected. One of the limiting factors is the lack of engineering techniques for distributed control systems design, simulation and verification. In this paper we analyze recently proposed techniques for distributed control systems development using an example of a simple transport system consisting of two CPS – smart conveyor belt and smart cylinder. In particular we consider the methods based on Control Interpreted Petri Nets (CIPN), Supervisory Control Theory (SCT) and IEC 61499 standard.**

***Index Terms*—Distributed Control; Industrial Control Systems; Cyber Physical Systems; Industry 4.0.**

I. Introduction

Industry 4.0 and introduction of Cyber Physical Systems (CPS) at manufacturing shop floor lead to significant changes in Industrial Control Systems (ICS) [1]. The demand for highly flexible automation and reconfigurable manufacturing induced by fluctuating market needs and high product variety on one [2], and the development of CPS based systems as the leading enabling technology on the other hand [1] represent the main drivers of these changes. CPS based smart devices with integrated computational and communication capabilities open up new possibilities in terms of ICS modularity, flexibility and reconfigurability. It is expected that with the full extent implementation of CPS at manufacturing shop floor the traditional automation hierarchy standardized through IEC 62264 will be replaced with truly distributed control systems where the control tasks will be carried out through interoperability of networked smart devices – Fig. 1. It is expected that all elements of automation hierarchy will remain, but in terms of functional hierarchy distributed over network without the pyramidal structure of the corresponding devices [3].

CPS are already readily employed at manufacturing shop floors in different automation tasks primary as smart sensors and actuators, and strict automation hierarchy is already broken down as there exist communication of different devices intra and over non-adjacent levels of automation pyramid. Nevertheless, as a rule, smart sensor and actuators are integrated in ICS in traditional manner – they are connected to the central controller (e.g., Programmable Logic Controller - PLC) that carries out the control task. In this way, computational capabilities of CPS, their modularity and ability to make manufacturing systems adaptable to new products are not fully exploited.

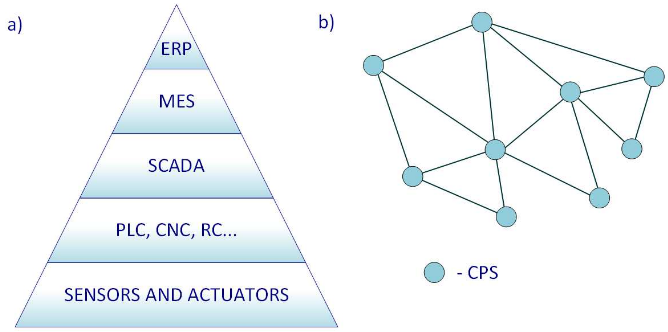


Fig. 1. Change of control paradigm in Industry 4.0: a) IEC 62264 automation hierarchy; b) Distributed control

There are several reasons for this. One is the inertia of control engineers community to implement new trends and their rationalle to keep the existing techniques for ICS design, that were practicaly tested and proven in a myraid of real-world examples. Yet, due to inability of traditional ICS to deal with high product variaty, there is a trend to perform a number of processes in manufacturing manually which represents a step backwards with respect to process sophistication; in this context, modularity of control system and distribution of control tasks are paramount for automation in Industry 4.0. The other reason is the lack of well-proven formaly-consequent engineering techniques for the desing of distributed control systems, i.e., for the distribution of control taks to smart devices [4]. Finally, when CPS are employed at manufacturing shop floor (regardless if centrally or distributed) cybersecurity related issues on communication links emerge.

Recently, a number of techniques for distribution of control tasks to smart devices and for dealing with cybersecurity within ICS have been proposed. Within this paper we will illustrate some of them using an example of simple system for parts transport that is presented in Section II. In particular, we will consider the method for control system distribution that is based on Control Interpreted Petri Nets (CIPN) that we proposed in [5], as well as the method based on Supervisory Control Theory [6] – Section III. Furthermore, in Section IV

Zivana Jakovljevic is with the University of Belgrade – Faculty of Mechanical Engineering, 16 Kraljice Marije, 11000 Belgrade, Serbia (email: zjakovljevic@mas.bg.ac.rs).

Dusan Nedeljkovic is with the University of Belgrade – Faculty of Mechanical Engineering, 16 Kraljice Marije, 11000 Belgrade, Serbia (email: dnedeljkovic@mas.bg.ac.rs).

we will show how the distributed control system can be simulated using IEC 61499 standard [7], and in Section V we will consider the possible effects of cyber-attacks on discrete event system using the methodology that we presented in [8]. Finally, Section VI gives some concluding remarks.

II. Description of the System and Control Task

The system that we consider in this paper is used for parts transport and consists of: 1) smart conveyer belt and 2) smart cylinder (Fig. 2). Conveyor belt transfers parts to position II and it is actuated by step motor M that is started/stopped using signal *m*. In addition, the belt contains sensor *s* that detects the presence of parts in removing position II, as well as start switch *st* that is used for the start of the system operation. Both sensors and actuator are connected to the belt's local controller denoted as LC_1. Smart cylinder, on the other hand, besides the actuator, contains a monostable 5/2 dual control valve controlled by signal *ap* (ap = 1 for cylinder advancing and ap = 0 for cylinder retracting), as well as limit switches $a1/a0$ for detecting final advanced/retracted position. Pneumatic and sensing devices are augmented with local controller LC_2 that controls the smart cylinder. The allocation of sensing and actuation signals to smart devices' local controllers is presented in Table I.

TABLE I
Sensors and Actuators Signals Mapping to Local Controllers

Device	Loc. Cont.	Signal	Description
Smart conveyer	LC_1	*m*	Motor actuation
		s	Sensor for part detection
		st	Start switch
Smart cylinder	LC_2	*ap*	Cylinder actuation
		$a1$	Advanced position sensor
		$a0$	Retracted position sensor

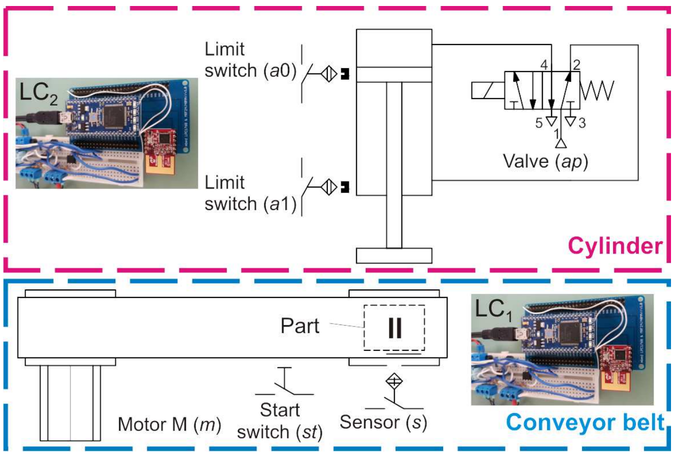


Fig. 2. Graphical representation of the system used in case study

The system should function as follows. After pressing the start switch, conveyor belt starts motion; when the part comes to the conveyor belt, it is transferred to position II where the sensor *s* is activated. After activation of signal *s*, conveyor belt stops and the cylinder removes the part from the belt (it advances and immediately retracts to home position). Once cylinder reaches the retracted position, the conveyor starts moving again to transfer new part to position II and the cycle repeats.

There are a number of different methods for formal description of the controller that would ensure the described system behavior when it is controlled using one (centralized) controller, e.g., PLC. Most frequently employed technique that represents *de facto* standard in ICS design are CIPN and the derived formalisms of GRAFCET (Graphe Fonctionnel de Commande des Étapes et Transitions - Functional Graph of Control by Steps and Transitions) and IEC 61131-3 Sequential Function Chart (SFC). CIPN represent bipartite graphs containing transitions represented by bars and places represented by circles [9]. Within CIPN each transition has associated condition as a Boolean function of sensory signals (although actuator signals can be used as well), whereas the actions which change the values of control system outputs are allocated to the CIPN places. The state of CIPN is represented by token(s) assigned to places, and during system evolution the token passes from previous to the next place when the transition between them fires, i.e., when the associated condition is true. Using CIPN formalism the desired performance of the system from Fig. 2 can be described using CIPN form Fig. 3; this CIPN can be easily transferred to SFC or other embedded devices programming languages and implemented in centralized PLC or some other device (e.g., microcontroller).

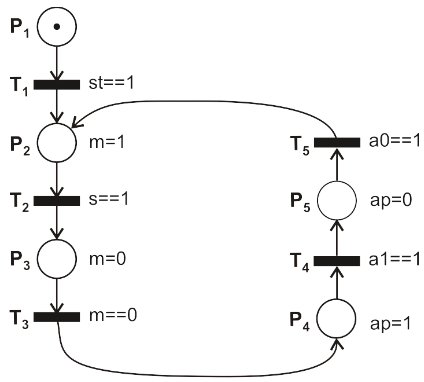


Fig. 3. CIPN describing the desired behavior of the system from Fig. 2.

III. Distribution of Control Tasks to Smart Devices

Opposite to centralized control systems design, generally, there is a lack of methods for the design of distributed controllers [10]. Existing approaches can be classified as *bottom-up* and *top-down*. Using *bottom-up* methods the behavior of each device within the control network is described using Petri nets [11], Supervisory Control Theory (SCT) [12], IEC 61499 [7] or similar formalisms, and the behavior of the system as a whole is obtained through their composition. These methods are characterized by low backwards compatibility since their implementation necessitates completely new mindset in system designers. Furthermore, they are usually based on trial and error and as such are time consuming and require significant verification [4]. *Top-down* approaches, on the other hand, describe the system as if centrally controlled and, using predefined methodology, distribute control tasks to smart devices. These methods are characterized by high backwards compatibility and can be easily embraced by engineers in everyday practice.

A *top-down* approach from [5] is based on CIPN. Global CIPN describing the behavior of the system as a whole and mapping of sensors and actuators to local controllers with direct access to these devices are at the input of this method. Following the set of rules that consider the allocation of sensing and actuation signals to transitions and places within global CIPN on one hand, and physical mapping of sensors and actuators to local controllers LC_i on the other, this method generates a separate $CIPN_i$ to be employed at each LC_i. Basically, this approach extracts from global CIPN into $CIPN_i$ the places that contain actuating signals allocated to LC_i along with preceding transitions, and introduces places with sending commands to compensate missing links. The details regarding the method can be found in [5].

Following this approach, for the CIPN from Fig. 3 and allocation of sensing and actuation signals from Table I, $\mathbf{CIPN_1}$ and $\mathbf{CIPN_2}$ presented in Fig. 4 are obtained. Within these CIPNs, the transitions and places that were extracted from global CIPN (Fig. 3) are denoted in parentheses. The actions associated to places within $\mathbf{CIPN_1}$ and $\mathbf{CIPN_2}$ contain only actuating signals mapped to corresponding devices (conveyor belt and cylinder, respectively). On the other hand the transitions can contain the signals allocated to the other device, as denoted red in T_4^1, T_{init}^2 and T_4^2. These signals are transferred between local controllers using selected communication protocol, and protocol agnostic sending commands are denoted green in corresponding places in Fig. 4. These commands are the result of the application of the procedure from [5].

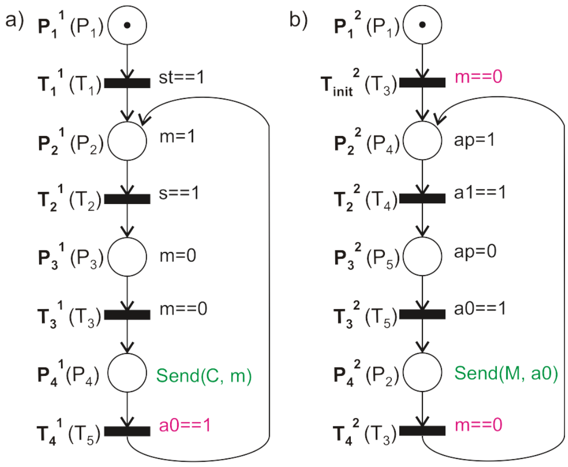


Fig. 4. CIPN representation of the system control distributed to smart devices: a) $\mathbf{CIPN_1}$ representing control tasks distributed to smart conveyer belt; b) $\mathbf{CIPN_2}$ representing control tasks distributed to smart cylinder;

For example, while at P_4^1 smart conveyor belt sends information about motor stopping to smart cylinder which receives it in transition T_{init}^2 or T_4^2 depending on the current state; in this way T_3 from Fig. 3 is performed. Similarly, while at P_4^2 smart cylinder sends information about getting to retracted position to smart conveyer belt, which the latter receives at transition T_4^1, delivering the T_5 from Fig. 3. Once CIPNs for local controllers are generated, their implementation is straightforward as in the case of centralized controllers.

An example of *bottom-up* approach for distribution of control tasks will be illustrated using SCT formalism [6]. The models of the local controllers – supervisors for smart devices can be presented by Finite State Machines (FSM) in which the transition between states occurs on the events that represent the change of certain (actuation or sensory) signal. FSMs representing the desired behavior of LC_1 and LC_2 are given in Fig. 5. In particular, S_B represents smart conveyer belt (Fig. 5a) and S_C (Fig. 5b) smart cylinder cyber part. Note that in smart conveyer belt FSM model motor activation signal (m = 1) is denoted by *mp*, whereas deactivation signal (m = 0) is denoted *mm*. Similarly, for smart cylinder ap = 1 is denoted by *ap* and ap = 0 by *am*. The signals that are transferred from a local controller are marked green, and the signals that are received are marked red.

As can be observed from the Fig. 5, the design of local controllers using this formalism is not straightforward and requires to simultaneously take care about the behavior of both devices and communication of signals between them which can be extremely tedious and error prone when larger number of local controllers is used. The equivalence of S_B and S_C to $\mathbf{CIPN_1}$ and $\mathbf{CIPN_2}$ from Fig. 4 can be observed.

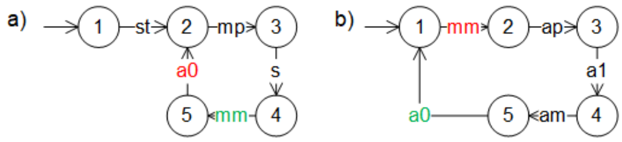


Fig. 5. Case study from Fig. 2 – FSMs representing: a) Conveyer belt local controller – S_B; b) Cylinder local controller – S_C;.

To verify that the system as a whole will have the desired behavior after implementation of the developed controllers, using SCT formalism, the physical part of the system should be modeled as well. FSMs from Fig. 6 - G_B and G_C represent physical parts of smart conveyer belt and smart cylinder, respectively.

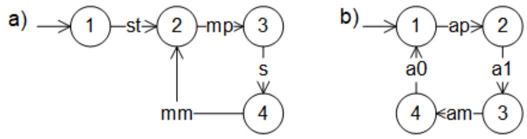


Fig. 6. Case study from Fig. 2 – FSMs representing: a) Conveyer belt physical device - G_B; b) Cylinder physical device - G_C.

Conjoint operation of all four FSMs from Figures 5 and 6 is presented by FSM G from Fig. 7, where x, y, z and v in state notation (x, y, z, v) represent the states of S_C, S_B, G_C and G_B, respectively. FSM G, that represents the behavior of the CPS from Fig. 2, is obtained using the following FSM operation:

$$G = (S_B \,||\, S_C) \times (G_B \,||\, G_C) \qquad (1)$$

where || denotes FSM parallel composition, and × product operator [6][1]. Comparing CIPN from Fig. 3 which represents the desired behavior of the system and FSM from Fig. 6 that represents the conjoint behavior of smart devices with distributed control tasks, the equivalence can be observed. Nevertheless, it should be noted that the implementation of SCT formalism assumes that CIPNs are not used during

[1] All automata operations are carried out using DESUMA software [13]

system representation, and in this paper we use it for the comparison only.

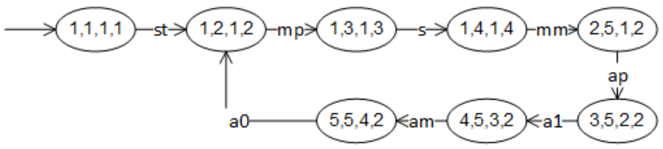


Fig. 7. Case study from Fig. 2 – behavior of the system represented using single FSM – *G*

IV. Modeling and Simulation of Distributed Control Systems using IEC 61499

Once the control task is distributed to smart devices' local controllers, it is beneficial to further verify it, preferably using simulation. IEC 61499: Function blocks [7] represents an international standard intended for distributed control system modeling and simulation. Using this standard the functionality of a system's hardware or software component is encapsulated using function blocks (FBs) that introduce object oriented paradigm into industrial control systems programming [14, 15]. FBs represent classes whose instances can be used for task execution in concrete applications. Using this formalism, CPS are modeled and simulated through interaction between their physical and cyber parts, each represented by separate FB. The blocks are integrated through real-time interaction that is modeled using events and data flows. The functionality of an FB is event driven and it is defined by its Execution Control Chart (ECC) that specifies the behavior of the modeled component when certain events in the system occur. They are in the form similar to CIPN.

IEC 61499 formalism models the behavior of the system through a network of FBs called application. Within application each smart device is introduced using corresponding FBs and multiple devices are interconnected using certain events and data. Fig. 8 represents IEC 61499 application for the system from Fig. 2[2]. Within this application function blocks ConveyCyber and ConveyPhys model cyber and physical parts of the smart conveyer, whereas CylCyber and CylPhys model corresponding parts of smart cylinder. Each FB contains head (upper part) that contains input (left side) and output (right side) events and body that contains input (left side) and output (right side) data. FB refreshes input/output data on the occurrence of the corresponding event, and their interconnections are modeled with connectors - red for event and blue for data flow.

Fig. 9 presents the ECCs for FBs modeling the behavior of the smart conveyor. At the system start, ConveyPhys transfers to S1, activates start switch (action *Start*) and issues **CNF** event to invoke the change of corresponding ConveyPhys output and ConveyCyber input; after that, it waits for **REQ** event and receipt of signal for motor activation (*mot* = *true*) from ConveyCyber when it passes to S2. Within S2 it activates sensor (action *Sensor_act*), issues **CNF** event to change corresponding input in ConveyCyber and waits for **REQ** events with the desired input data from ConveyCyber to return to S1.

[2] For IEC 61499 modeling and simulation 4diac software [16] is used.

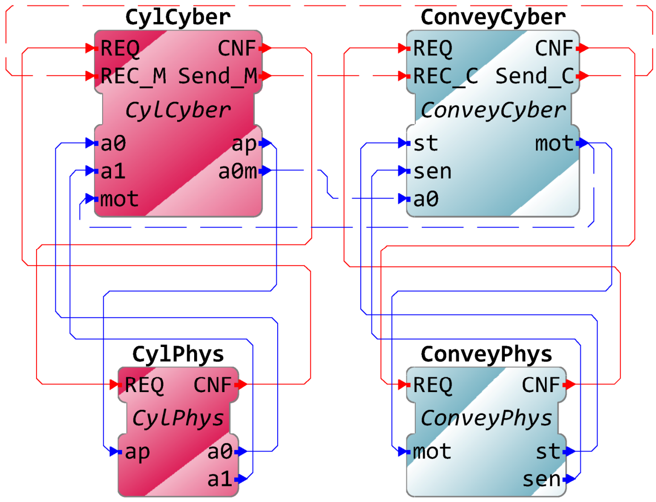
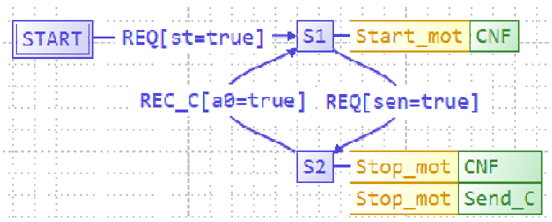
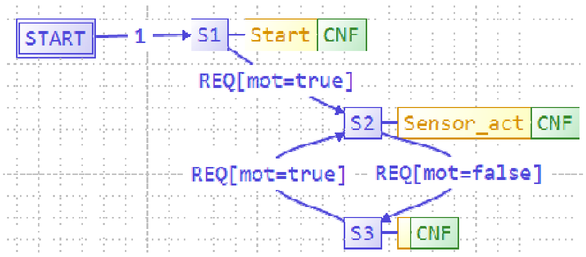


Fig. 8. IEC 61499 application that models the behavior of the system from Fig. 2.

a)



b)



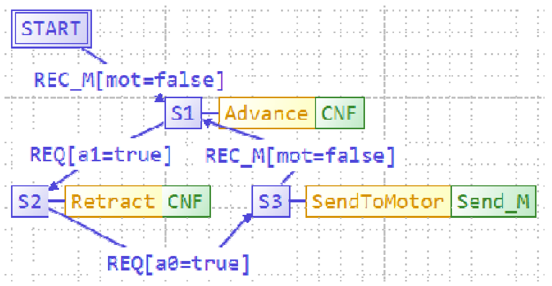
c)

ConveyCyber actions	
Start_mot	mot:=**true**;
Stop_mot	mot:=**false**;
ConveyPhys actions	
Start	st:=**true**;
Sensor_act	sen:=**true**;

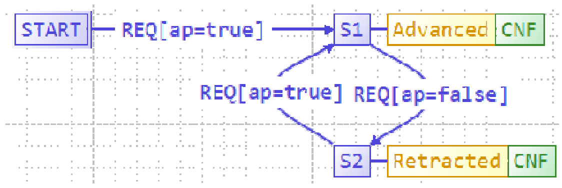
Fig. 9. ECCs for conveyer belt: a) cyber part of the conveyer belt - ConveyCyber ; b) physical part of the conveyer belt ConveyPhys; c) actions definition.

Simultaneously, at the beginning, ConveyCyber waits for **REQ** event (**CNF** event from ConveyPhys – Fig. 8) and the value *st* = *true* from ConveyPhys; on the receipt of this data it transfers to S1 where it carries out *Start_mot* action and invokes **CNF** event to inform the ConveyPhys (which is at that moment in S1 or S3) about signal change. Afterwards it waits in S1 information from ConveyPhys that sensor is activated to transfer to S2. In S2 it invokes the action for motor stopping and informs ConveyPhys and CylCyber about this change using events **CNF** and **Send_C**. Finally, after receipt of information that cylinder removed the part from the belt from CylCyber through event **REC_C** and data $a0$ = *true*, it returns to S1.

a)



b)



c)

CylCyber actions	
Advance	ap:=**true**;
SendToMotor	a0m:=**true**;
Retract	ap:=**false**;
CylPhys actions	
Advanced	a0:=**false**; a1:=**true**;
Retracted	a0:=**true**; a1:=**false**;

Fig. 10. ECCs for cylinder: a) cyber part of the cylinder - CylCyber ; b) physical part of the cylinder CylPhys; c) actions definition.

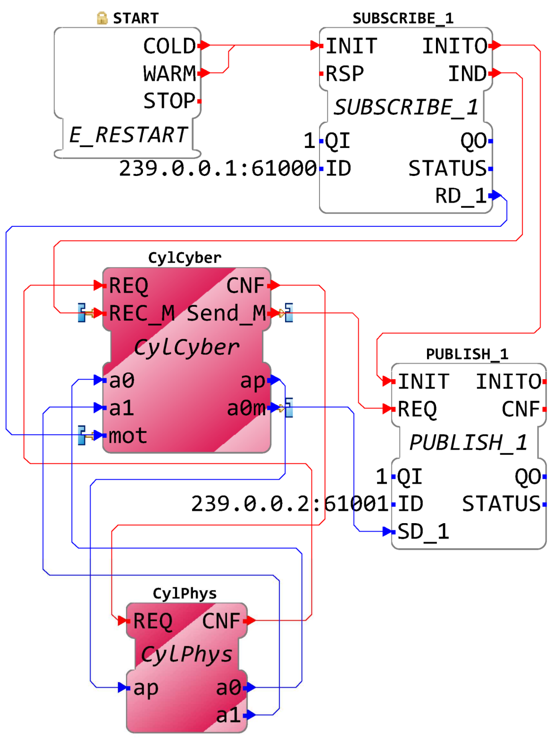


Fig. 11. Cylinder resource with introduced Publish and Subscribe function blocks for simulation of the system performance.

Similarly (Fig. 10), at the begining CylCyber waits for the information from ConveyCyber that motor stopped (**REC_M** event that is connected to **Send_C** from ConveyCyber – Fig. 8), moves to S1, invokes cylinder advancing (action *Advance*), and upon receipt of the event and data referring to the completion of CylPhys advancing, issues commands for its retraction. When it gets the information that CylPhys retracted, CylCyber issues **Send_M** event to inform CylCyber that motor can start motion, and waits for the new signal from ConveyCyber that motor stopped (**REC_M** event connected to **Send_C**) to return to S1 and prepare for the new cycle. Finally, CylPhys moves between states corresponding to advanced and retracted positions in accordance with the events and data received from CylCyber.

Comparing these ECCs with CIPNs and FSMs from Figures 4, 5 and 6, the equivalence can be observed.

To simulate the behavior of the control system defined by application (Fig. 8) function blocks are allocated to different resources and the communication is introduced through publish and subscribe FBs as presented in Fig. 11 using an example of smart cylinder.

V. Security Related Issues

As can be observed from previous sections, the performance of distributed control system is communication intensive, thus openings up the space for malicious cyber-attacks by various adversaries and raising cyber security related issues. A number of different attack scenarios can be identified and in case of discrete event systems such as one at hand two kinds of attacks can be singled out: 1) event removal and 2) event insertion. The adversaries can issue these attacks in various combinations always having the common goal to remain stealthy and to have negative effect on the system performance. Considering the possible consequences of cyber-attacks on ICS that can be not only of economic nature but also safety related (catastrophic damages of the equipment or even threats to human health and lives), it is crucial to analyze possible attack points, scenarios and consequences at early phases of the system design. Within this paper we will briefly illustrate an approach for modeling attacks scenarios using SCT based methodology that we proposed in [8]. We will consider the examples of removal and insertion attacks on $a0$ signal transmitted from smart cylinder to conveyor belt.

Following $a0$ removal attack denoted by $a0r$, smart conveyor belt cyber part will remain in the state at which it waits for $a0$ to progress with functioning, whereas smart cylinder cyber and physical parts will continue functioning as if attack did not occur. Described evolution of the system can be modeled using FSMs under attack $S_B{}^r$, $S_C{}^r$ and $G_C{}^r$ presented in Fig. 12a-c. In these automata the evolution of the system elements under attack is denoted in red lines. The details of the model generation can be found in [8], and the performance of the system as a whole is represented by automaton:

$$G^r = \left(S_B^r \,||\, S_C^r\right) \times \left(G_B \,||\, G_C^r\right) \qquad (2)$$

which is presented in Fig. 12d. From this figure it can be observed that on the $a0$ removal attack the system stops (it enters the state marked red) and no damage is expected.

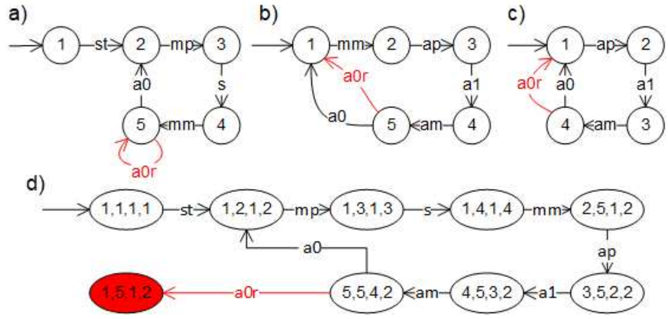


Fig. 12. Model of the system under $a0$ removal attack: a) S_B^r – model of conveyor cyber part; b) S_C^r – model of cylinder cyber part; c) G_C^r – model of cylinder physical part; d) model of the whole system behavior under attack.

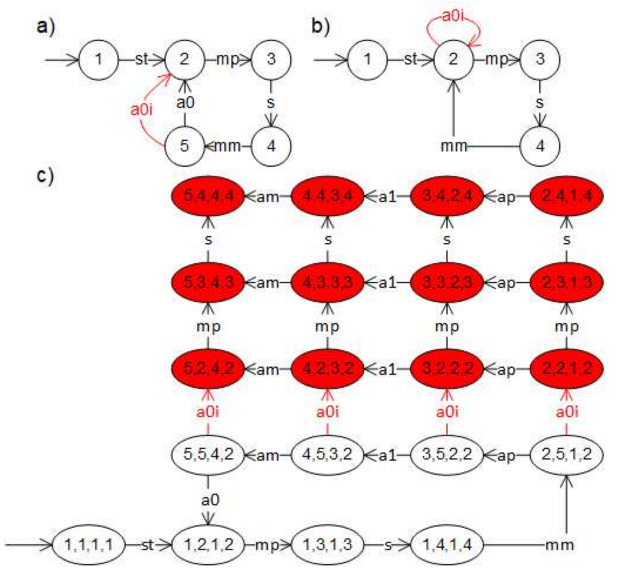


Fig. 13. Model of the system under $a0$ insertion attack: a) S_B^i – model of conveyor cyber part; b) G_B^i – model of conveyor physical part; c) model of the whole system behavior under attack.

When insertion attack is considered, it should be noted that to remain stealthy, the adversary can issue the $a0$ insertion attack only while S_B is in state 5 [8]. Following insertion attack $a0i$, the conveyer belt supervisor will transfer from state 5 to state 2 as if real $a0$ was received as presented in red line in FSM S_B^i in Fig. 13a. Simultaneously, conveyor physical part will remain in the state 2 as given in automaton G_B^i from Fig. 13b (the details regarding the formalism for generation of these automata can be found in [8]). The behavior of the whole system under $a0$ insertion attack is modeled through:

$$G^i = \left(S_B^i \,||\, S_C\right) \times \left(G_B^i \,||\, G_C\right) \quad (3)$$

presented in Fig. 13c where the possible evolutions of the system after attack are denoted using red states. These scenarios include starting the motor before the part is removed from the conveyer belt and can lead to the falling of the part from the transporter and its damage.

VI. Conclusion

Within this paper we have summarized and illustrated using a case study our recent research results in the area of CPS based distributed control systems design, verification and cyber protection. In particular, we have presented the application of the *top-down* approach for distributed control system design that is based on CIPN and that is characterized by excellent backwards compatibility. The convenience of this technique is supported through illustrating the application of a *bottom-up* technique based on SCT that requires completely new approach to ICS modeling. Furthermore, we have shown how the developed distributed control system can be simulated using standard IEC 61499. Finally, since the deployment of CPS and IoT at manufacturing shop floor leads to significant cyber security issues, we have shown how the SCT based framework can be applied for the analysis of cyber threats in our case study.

Acknowledgment

This research was supported by the Science Fund of the Republic of Serbia, grant No. 6523109, AI-MISSION 4.0

References

[1] H. Kagermann, W. Wahlster, and J. Helbig, "Recommendations for Implementing Strategic Initiative INDUSTRIE 4.0," Acatech, Germany, 2013 [Online]. Available: http://www.acatech.de
[2] H. ElMaraghy, "Smart changeable manufacturing systems." *Proc Manuf*, vol. 28, pp. 3-9, 2019.
[3] Reference Architecture Model Industrie 4.0 (RAMI4.0), 2015. [Online]. Available: fhttp://www.zvei.org
[4] J. Otto, B. Vogel-Heuser, and O. Niggemann, "Automatic parameter estimation for reusable software components of modular and reconfigurable cyber-physical production systems in the domain of discrete manufacturing," IEEE *Trans. Ind. Informat.*, vol. 14, no. 1, pp. 275–282, Jan. 2018.
[5] Z. Jakovljevic, V. Lesi, S. Mitrovic, and M. Pajic, "Distributing sequential control for manufacturing automation systems," *IEEE Trans. Contr Syst Tech.*, vol. 28, no. 4, pp. 1586-1594, July 2020.
[6] P. Ramadge and W. Murray Wonham, "The control of discrete event systems," *Proc. IEEE*, vol. 77, no. 1, pp. 81–98, 1989.
[7] *Function blocks – Part 1: Architecture*, IEC 61499-1:2012.
[8] Z. Jakovljevic, V. Lesi, and M. Pajic, "Attacks on Distributed Sequential Control in Manufacturing Automation," *IEEE Trans. Ind. Informat.*, vol. 17, no. 2, pp. 775-786, Feb. 2021.
[9] R. David and H. Alla, *Discrete, Continuous, and Hybrid Petri Nets*, 2nd ed. Berlin, Germany: Springer, 2010.
[10] R. Vrabič, D. Kozjek, A. Malus, V. Zaletelj, and P. Butala, "Distributed control with rationally bounded agents in cyber-physical production systems," *CIRP Ann.*, vol. 67, no. 1, pp. 507–510, 2018.
[11] L. Wang, C. Mahulea, and M. Silva, "Distributed model predictive control of timed continuous Petri nets," in Proc. IEEE Conf. Decis. Control, pp. 6317–6322, Dec. 2013.
[12] M. Wakaiki, P. Tabuada, and J. P. Hespanha, "Supervisory control of discrete-event systems under attacks," Dynamic Games and Appl., vol. 9, pp 965-983, Dec. 2019.
[13] L. Ricker, S. Lafortune, and S. Genc, "Desuma: A tool integrating giddes and umdes," in Proceedings - Eighth International Workshop on Discrete Event Systems, WODES 2006, pp. 392–393, 2006.
[14] J. H. Christensen, T. Strasser, A. Valentini, V. Vyatkin, and A. Zoitl, "The IEC 61499 Function Block Standard: Software Tools and Runtime Platforms," in ISA Automation Week, 2012.
[15] V. Vyatkin, *IEC 61499 Function Blocks for Embedded and Distributed Control Systems Design*, ISA, 2012.
[16] https://www.eclipse.org/4diac/ Accessed: 2020-04-24

Application of the Angular Dependency of the Zero Moment Point

Tilen Brecelj[1] and Tadej Petrič[2]

***Abstract*— In this paper the widely used stability parameter called the zero moment point (ZMP) is redefined as the angle around the center of mass (COM) of the investigated system. With this redefinition the ZMP is expressed in a more general way which enables its application in a wider range of situations. The angular definition of the ZMP was validated with motion measurements of a person performing different movements recorded with the OptiTrack camera system. The skeleton of the filmed person was reconstructed with the Motive software and a body model was used to reconstruct its COM, which was further used to calculate the ZMP. The stability analysis of the recorded motion measurements presented in this article shows on real-world examples of human motion that the angular redefinition of the ZMP provides a general, reliable and simple-to-apply way of determining the stability of a system.**

I. INTRODUCTION

Since the beginnings of motion control researchers have investigated stability conditions and defined different parameters that reveal whether a systems is stable or not and enable the calculation of possible motions in different situations. The research in the field of robot stability has deepened with the development of robotic systems capable of performing complex motion tasks such as for example humanoid robots performing human-like motion. Achieving stability of humanoid robots can be due to their relatively small feet in comparison to their relatively large body sizes very challenging. This is why also simple tasks such as for example walking, demand for constant stability verification and prediction in order to enable their accomplishment [1], [2], [3], [4]. But as the progress in this field is advancing with an incredible speed, robots became recently also capable of running, jumping and even skiing [5], [6], [7].

A system is capable of performing the desired tasks only if the forces, acting from the support polygon, allow it, which means that they make the system dynamically or statically stable. If the system is supported only by the forces acting from the ground, the support polygon lies within the boundaries of the contact of the system with the ground. In the case of a humanoid supported only by its feet the support polygon extends from its heel to its toes and between the outer edges of its left and its right foot.

One of the parameters defining the system's stability is the zero moment point (ZMP), defined as the location on the ground, where all the forces and torques, acting on the system, can be replaced by only one force [8], [9]. If the ZMP lyes within the support polygon it coincides with the center of pressure and in this case the system is stable. On the other hand, if the ZMP lies outside the support polygon it can not exist as there are no support mechanism outside the support polygon and therefore it is usually refereed to as a fictitious ZMP. In this case the system is not stable and if the system is a humanoid it will flip either over its toes or its heel.

But different systems may have different support mechanism that do not necessarily act on the ground. A humanoid may for example support itself with his arms, that may apply support forces at different heights, or with its bottom, in the case when it is sited. In such scenarios the support polygon does not lie only on the ground but it extends to different heights, as shown in Fig. 1, and therefore the standard

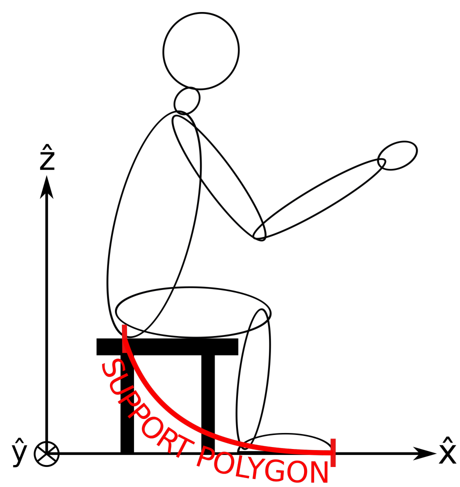


Fig. 1: The support polygon of a humanoid sited on a bench.

definition of the support polygon and the ZMP located on the ground can not be used. This is why in this article the ZMP is expressed in a more general way as the angle around the center of mass (COM) of the investigated system. Such definition can be applied to all systems, also those supported by different mechanisms at different heights, even below the ground or above the system itself.

In this article the stability of a humanoid, which can be either a person or a humanoid robot, is described, but the derived stability conditions can be applied to any system.

II. THE ZERO MOMENT POINT

A. The Standard Cartesian Definition

One of the most widely used stability parameters, the ZMP, can be expressed from the torque balance equation that takes into account the gravitational force acting on each segment of the humanoid, the forces accelerating each segment of the

[1]Tilen Brecelj and [2]Tadej Petrič are with the Jožef Stefan Institute, Jamova c. 39, 1000 Ljubljana, Slovenia {tilen.brecelj, tadej.petric}@ijs.si

humanoid in the desired direction, along with the torques produced by the rotations of each humanoid segment and all the external torques and forces present. By the definition the horizontal torques in the ZMP must be zero and only torques in the vertical direction can exist, but they are in usual circumstances balanced by the friction forces [8], [9].

The coordinate system defining the space in which the humanoid is positioned is oriented such that the sagittal plane of the humanoid lies in the $x - z$ plane, the lateral plane of the humanoid lies in the $y - z$ plane, while the ground lies in the $x - y$ plane, as shown in Fig. 1. This way the x and y components of the torque in the ZMP must be equal to zero, while its z component can be non-zero. For a thorough explanation of the torque balance equation and the derivation of the ZMP see [10].

The torque balance equation is a general equation that takes into account multiple effects that do in numerous circumstances not exist or can be neglected. This is why the position of the ZMP on the ground is usually obtained with some simplified models, one of them being the linear inverted pendulum model [11]. The latter assumes that the investigated system is symmetric with respect to the sagittal plane, it can be approximated by its COM, there are no external forces and torques present, the system does not rotate around any axis and that its COM lies at $y = 0$ and moves only in the x direction. In this case the location on the ground where the ZMP lies can be obtained as

$$x_{\text{zmp}} = x_{\text{com}} + \frac{z_{\text{com}}}{g} \ddot{x}_{\text{com}} \,, \tag{1}$$

where x_{com} and z_{com} are the x nad z coordinates of the COM of the investigated system, respectively, $\ddot{x}_{\text{com}}$ is its acceleration in the x direction, while g is the gravitational acceleration.

But the Cartesian definition of the ZMP may become useless in practice if the ZMP is located too far away from the investigated system, as the system may not have suitable mechanisms to support itself at such distant locations. Furthermore, if the system uses other support mechanisms that are not applying support forces on the ground, which may be in the case of a humanoid its arms and hands, the support polygon does not extend only on the ground and therefore its stability can not be verified using only the ground location of the ZMP. This is why a more general expression of the ZMP was developed.

B. The Angular Definition

From the torque balance equation it can be shown that taking into account the same assumptions as for the derivation of the linear inverted pendulum model, with the exceptions that now the ZMP does not need to lie on the ground any more and that the COM of the system can be accelerated also in the z direction, the ZMP can lie anywhere on the ZMP line defined as

$$z_{\text{zmp}} = \tan^{-1}(\varphi_{\text{zmp}})(x_{\text{com}} - x_{\text{zmp}}) + z_{\text{com}} \,, \tag{2}$$

where z_{zmp} and x_{zmp} are the z and x coordinates on the ZMP line, respectively, φ_{zmp} is the ZMP angle defined as

$$\varphi_{\text{zmp}} = -\arctan 2\big(\ddot{x}_{\text{com}}, (\ddot{z}_{\text{com}} - g)\big) \tag{3}$$

and $\ddot{z}_{\text{com}}$ is the vertical acceleration of the COM. φ_{zmp} is located between the vertical line, passing through the COM of the investigated system and the ZMP line, which is in the case, when the assumptions made for the derivation of (2) hold true, passing through the COM of the investigated system, as shown in Fig. 2. With the angular redefinition

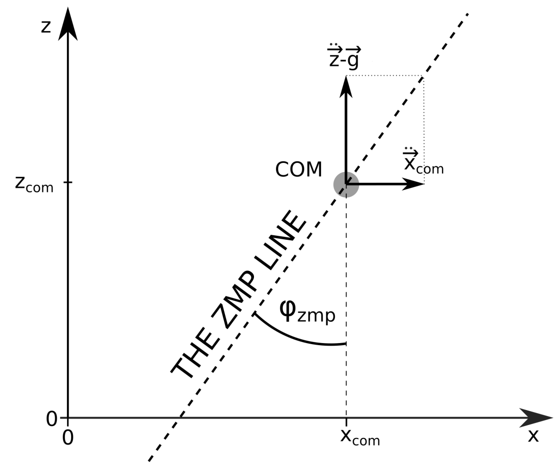


Fig. 2: The angular definition of the ZMP. φ_{zmp} is the angle between the ZMP line and the vertical line, passing through the COM of the investigated system.

the ZMP does not need to lie on the ground any more but it can lie at any height also below the ground, where $z_{\text{zmp}} < 0$, anywhere between the ground and the COM of the investigated system, where $0 \leq z_{\text{zmp}} \leq z_{\text{com}}$, or above the COM, where $z_{\text{zmp}} > z_{\text{com}}$. For a detailed derivation of the angular definition of the ZMP see [10]. To obtain the ZMP angle in the lateral $(y - z)$ plane of the humanoid, the COM coordinates and accelerations in the x direction from (2) and (3) must be substituted with the corresponding values in the y direction.

The stability condition stating that the ZMP must lie within the support polygon can be expressed with the ZMP angle and the angles of the edges of the support polygon in the following way. For the sagittal plane of the humanoid, which is supported by only one foot, or by both feet, positioned in parallel one next to each other, this condition can be expressed as

$$\varphi_{\text{heel}} \leq \varphi_{\text{zmp}} \leq \varphi_{\text{toes}} \,, \tag{4}$$

where

$$\varphi_{\text{heel, toes}} = \arctan\left(\frac{x_{\text{heel, toes}} - x_{\text{com}}}{z_{\text{com}}}\right), \tag{5}$$

as shown in Fig. 3 If the humanoid is supported on the ground with both feet, that are not parallel one to the other, φ_{heel} refers to the heel of the back foot, while φ_{toes} refers to the toes of the front foot. If, on the other hand, the humanoid

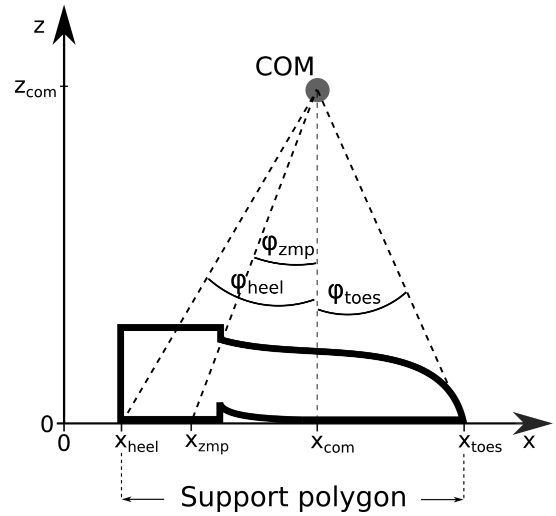


Fig. 3: The angles of the edges of the heel and the toes of the foot of the humanoid with respect to its COM, as defined by (5).

is oriented in the opposite direction, φ_{heel} and φ_{toes} must be interchanged. If the condition (4) does not hold true, the humanoid is not stable and it will flip over its toes or its heel.

III. Validation of the Angular Definition of the ZMP

The angular definition of the ZMP was validated on two measurements of human motion filmed with a set of 13 OptiTrack cameras [12], emitting infrared light and detecting its reflection from 37 reflective markers, positioned on the filmed person. The markers were placed on predefined positions on the human body in order to enable Motive 2.2.0 [13] (the OptiTrack optical motion capture software) to reconstruct the human skeleton. Knowing the approximate positions of the human joints, estimated by the Motive software, and using a human body model [14], [15], [16], which enables the reconstruction of the sizes and masses of each body segment and the distances of their corresponding centers of mass (COMs) from the adjacent joints, the reconstruction of the total COM of the filmed person was possible. The ZMP line and the ZMP angle from (2) and (3), respectively, were then obtained from the reconstructed location and acceleration of the total COM of the person for each recorded frame.

In Fig. 4 you can see the frame sequence outtake from the first filmed motion. The top row shows the lateral plane (front view), while the bottom row shows the sagittal plane (side view) of the filmed person. The frames positioned one above the other were obtained at the same time and therefore represent the same capture. On the first two captures, obtained at times 2.7 s and 4.1 s, the filmed person was swinging from his right side to his left side, on the second two captures, obtained at times 20.6 s and 24.9 s, the person was bowing forth and back, while at the last capture, obtained at time 37.6 s, the person was standing on only one leg, bowing forth, with the other leg lifted up in the air and with his hands extended sideways.

The y and the x coordinates of the ZMP and the edges of the support polygon of the recorded person during the first motion measurement, presented in Fig. 4, where the person was supported only by his feet, are shown in Fig. 5 as a function of time. On the plot on top it can be seen that in the y direction the person was unstable only for small amounts of time during the measurement, when the ZMP is located outside the support polygon. During this measurement the person was not standing on both feet all the time but was also stepping to only one foot while swinging from one side to the other. This can be seen as a sudden narrowing of the support polygon in the y direction, such as for example at time 1.2 s, when the person started to stand only on his left foot, and a sudden widening of the support polygon, when the person was supported again by both feet, such as for example at time 2.0 s. The deviation of the ZMP outside the support polygon occurred during the swinging at around 1.9 s and 15.6 s, when the person was supported only by his left and only by his right foot, respectively, just before he landed to the other foot. This instability arised because just before the landing on the second foot, the person was falling down towards the ground and could not control his motion. Another instability can be seen at time around 14.1 s when the person was standing only on his right foot and was caused by fast and sudden movements of the person while catching balance. As the ZMP was calculated only from the reconstructed position and acceleration of the total COM of the body, the torques produced by the rotations of different body links were not taken into account, which can for fast and sudden movements influence the location of the ZMP. On the other hand, in the x direction the person was stable all the time, as the ZMP location was within the boundaries of the support polygon throughout the whole measurement, as shown in the bottom plot of Fig. 5.

Fig. 6 shows the same support polygon as Fig. 5, but with the ZMP and the edges of the support polygon expressed as angles around the COM of the measured person, defined by (3) and (5), respectively. The angular results are similar to the results expressed with Cartesian coordinates on the ground, but the lines representing the edges of the angular support polygon are more curved than the corresponding lines in the Cartesian coordinate system, as the angular results are not obtained relative to a fixed coordinate system but relative to the COM of the moving person. The angular values of the ZMP are, on the other hand, subjected to smaller variations in time than the corresponding Cartesian values.

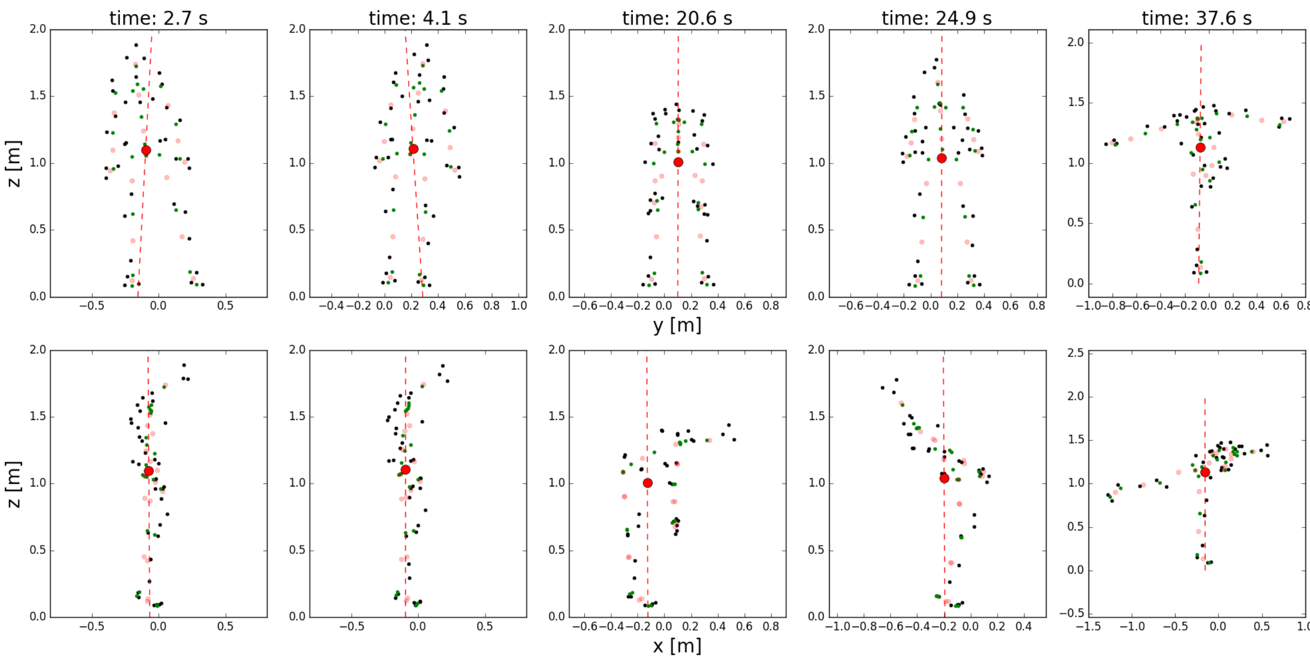


Fig. 4: Outtakes of the recorded and reconstructed data of the first filmed motion. The black dots represent the locations of the markers positioned on the filmed person, the green dots represent the reconstructed joint positions by the Motive software, the pale red dots represent the reconstructed locations of the COMs of each body link, the big red dot represents the location of the total COM of the filmed person and the red dotted line is the ZMP line. For the explanation of the body postures in each frame see the text.

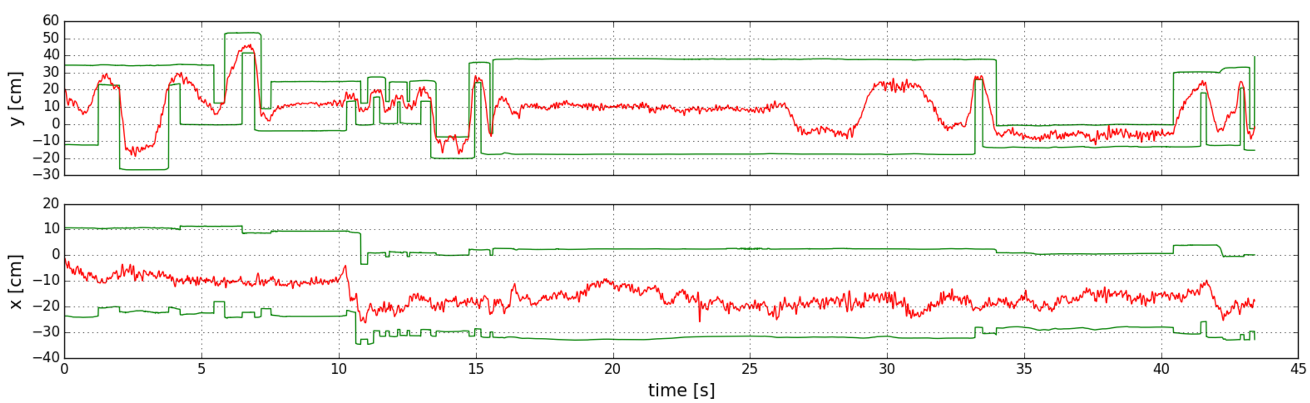


Fig. 5: The y (top) and x (bottom) coordinates of the ZMP (red lines) and the edges of the support polygon (green lines) for the first motion measurement expressed on the ground.

In the second motion filmed with the OptiTrack cameras the person is sitting down on a bench and standing up. Fig. 7 shows the frame sequence outtake of this motion.

In the first capture at time 0.2 s the person was stepping in front of the bench, in the second capture at 1.6 s the person was sitting down with his hands positioned on the bench, in the third capture at time 5.4 s the person was seated and in the forth and fifth captures at 8.2 s and 8.8 s, respectively, the person was standing up.

As throughout this motion the person was not supported only by his feet on the ground but also by his hands and bottom on the bench, the support polygon does not extend only on the ground but also on the bench. This is why the standard Cartesian definition of the ZMP and the support polygon on the ground can not be used but the angular redefinition of these quantities must be applied. Fig. 8 shows the angles of the ZMP and the edges of the support polygon obtained with (3) and (5), respectively. When the person was supported also by his hands and his bottom on the bench, the x coordinate of his hell from (5) was substituted with the x coordinate of his hand or bottom, while the height

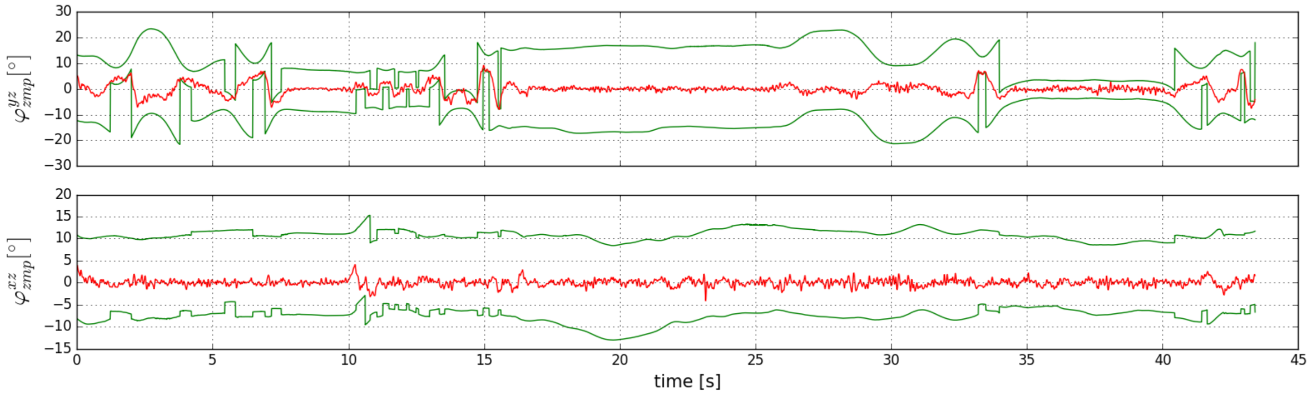


Fig. 6: The angles of the ZMP (red lines) and the edges of the support polygon (green lines) in the lateral (top) and the sagittal (bottom) planes of the measured person, for the first motion measurement

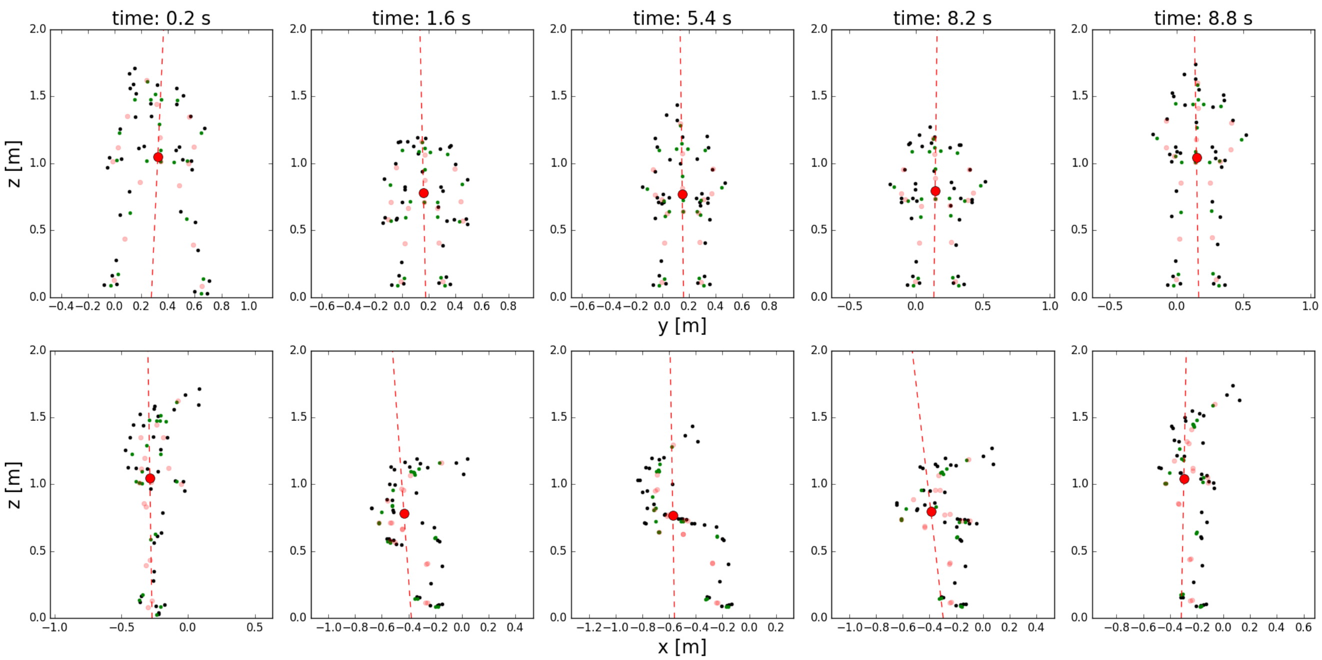


Fig. 7: Outtakes of the recorded and reconstructed data of the second filmed motion of sitting and standing. For the explanation of the meaning of different symbols see the text under Fig. 4. For the explanation of the body postures in each frame see the text.

of the COM was recalculated relatively to the bench height. In this measurement the ZMP was always within the support polygon which means that the person was stable all the time. In the beginning of the filming when the person started to sit down, he was supported only by his feet. At time 1.3 s he placed his fingers and at time 1.5 s also his palms on the bench, which can be seen as a widening of the support polygon angles in both the lateral and the sagittal planes, as the support polygon extended from his feet to the locations on the bench, where he was supported. At time 1.7 s the person sat on the bench and lifted up his arms

from the bench, which can be seen as a narrowing of the support polygon in the lateral plane and widening of the support polygon in the sagittal plane. The person was sited till the time 7.8 s, when he started to stand up which caused the narrowing of the support polygon in the sagittal plane, as the outermost location on his thighs where he was still in contact with the bench was approaching the edge of the bench above his feet. In the moment when he detached from the bench and was supported only by his feet at 8.1 s, the ZMP moved within the support polygon limited by the edges of his feet. Immediately after he was supported only by his feet the ZMP angle was in the sagittal plane close to the angle of his heel, but when he straightened up, the ZMP angle moved approximately in the middle between the angles

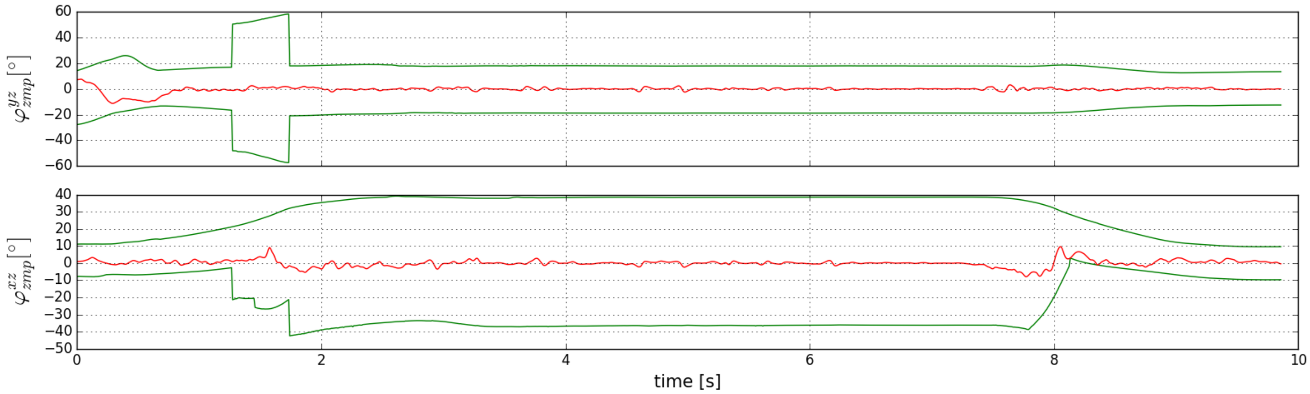


Fig. 8: The angles of the ZMP (red lines) and the edges of the support polygon (green lines) in the lateral (top) and the sagittal (bottom) planes of the measured person, for the second measurement of sitting and standing.

of his heels and his toes.

IV. CONCLUSIONS

The standard Cartesian definition of the ZMP on the ground is not suitable for the stability analysis in the situations when for example the ZMP lies at large distances from the investigated systems or if the system is supported at different heights. But in all these scenarios the angular definition of the ZMP can be used. If the system has support mechanisms that can act at different heights, such as for example a humanoid with arms, the situations when the ZMP would lie at far distances from the humanoid on the ground can be easily handled if the humanoid is supported at angles that embed the ZMP angle and if the friction forces and the forces in his joints allow for the satisfaction of the torque balance equation.

In this article it is shown how the stability of a person can be examined in the case when the person is supported only by his feet on the ground and in the case when the person is supported at different heights. In the first case both definitions of the ZMP and the support polygon were used, the one that defines these quantities in the Cartesian coordinates on the ground and the one that defines them as angles around the COM of the measured person. But in the second case, when the person was sitting on a bench and standing up, he was supported at different heights and the standard Cartesian definition of the ZMP and the support polygon on the ground could not be used and therefore our angular redefinition was applied. This way the stability of the person could be monitored in all the situations, no matter where the person was supported and whether the support mechanisms were only his feet or also his hands and his bottom. But the angular definitions of the ZMP and the support polygon are general and can therefore be applied to any system.

ACKNOWLEDGMENT

This work was supported by the Slovenian Research Agency grant N2-0153.

REFERENCES

[1] T. Petrič, L. Žlajpah, G. Garofalo, and C. Ott, "Walking control using adaptive oscillators combined with dynamic movement primitives," 01 2013.
[2] T. Petrič, A. Gams, J. Babič, and L. Žlajpah, "Reflexive stability control framework for humanoid robots," *Autonomous Robots*, vol. 34, 05 2013.
[3] A. J. Ijspeert, "Central pattern generators for locomotion control in animals and robots: A review," *Neural Networks*, vol. 21, no. 4, pp. 642 – 653, 2008, robotics and Neuroscience.
[4] L. Righetti and A. Ijspeert, "Programmable central pattern generators: An application to biped locomotion control," vol. 2006, 06 2006, pp. 1585 – 1590.
[5] J. Babič and J. Lenarčič, "Optimization of biarticular gastrocnemius muscle in humanoid jumping robot simulation," *Int. J. Humanoid Robotics*, vol. 3, pp. 219–234, 2006.
[6] L. Lahajnar, A. Kos, and B. Nemec, "Skiing robot - design, control, and navigation in unstructured environment," *Robotica*, vol. 27, pp. 567–577, 07 2009.
[7] T. Petrič, A. Gams, J. Babič, and L. Žlajpah, "Reflexive stability control framework for humanoid robots," *Autonomous Robots*, vol. 34, 05 2013.
[8] M. Vukobratovič and D. Juričić, "Contribution to the synthesis of biped gait." *IEEE transactions on bio-medical engineering*, vol. 16 1, pp. 1–6, 1969.
[9] M. Vukobratovič and B. Borovac, "Zero-moment point - thirty five years of its life." *I. J. Humanoid Robotics*, vol. 1, pp. 157–173, 03 2004.
[10] T. Brecelj and T. Petrič, "Angular dependency of the zero moment point," in *Advances in Service and Industrial Robotics*, S. Zeghloul, M. A. Laribi, and J. Sandoval, Eds. Cham: Springer International Publishing, 2021, pp. 135–144.
[11] S. Kajita and K. Tani, "Study of dynamic biped locomotion on rugged terrain-derivation and application of the linear inverted pendulum mode," in *Proceedings. 1991 IEEE International Conference on Robotics and Automation*, 1991, pp. 1405–1411 vol.2.
[12] Optitrack motion capture and 3d tracking system. https://optitrack.com/.
[13] Motive optical motion capture software. https://optitrack.com/software/motive/.
[14] P. de Leva, "Adjustments to zatsiorsky-seluyanov's segment inertia parameters," *Journal of Biomechanics*, vol. 29, no. 9, pp. 1223–1230, 1996. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0021929095001786
[15] Y. T. E. Todorov and T. Erez, *MuJoCo: Modeling and Simulation of Multi-Joint Dynamics with Contact*, 0th ed., 2013.
[16] V. Zatsiorsky and V. Seluyanov, in *Biomechanics VIII-B*. Human Kinetics, 1983, p. 1152–1159.

ТЕЛЕКОМУНИКАЦИЈЕ
/
TELECOMMUNICATIONS

(TE/TEI)

Eksperimentalna karakterizacija turbulencije u bežičnom optičkom kanalu

Dejan Milić, *Member, IEEE*, Jelena Anastasov, *Student Member, IEEE*, Daniela Milović i Nenad Milošević, *Member, IEEE*

***Apstrakt*— U ovom radu prikazani su eksperimentalni rezultati dobijeni u laboratorijski kontrolisanim uslovima za karakterizaciju pojave turbulencije u slobodnom prostoru pri optičkom prenosu (FSO - free space optics). Prikupljeni podaci korišćeni su za statističku karakterizaciju funkcije gustine verovatnoće (pdf - probability density function) fluktuacija optičkog signala. U literaturi postoji niz empirijskih modela za različite režime turbulencije. Dobijeni pdf je upoređivan sa odabranim modelima iz literature u cilju definisanja statističkih osobina realnog turbulentnog kanala što je značajno za proračun performansi FSO sistema. Analiza je pokazala da se rezultati dobijeni pod datim eksperimentalnim uslovima dobro poklapaju sa eksponencijalno - Vejbulovim modelom turbulencije.**

***Ključne reči*—funkcija gustine verovatnoće; FSO link; turbulencija; scintilacija.**

I. Uvod

Optičke komunikacije kod kojih se optički signali prostiru kroz slobodni prostor (FSO - free space optical communication) nude pristup optičkom delu spektra, rešavajući tako probleme sa zagušenjem radio spektra [1], [2]. Pored toga, FSO sistemi su značajni kao odgovor na sve veću potrebu za brzom bežičnom komunikacijom. Danas, FSO sistemi se upotrebljavaju u lokalnim i metro mrežama, u komunikaciji između zemaljskih stanica i satelita [3]. Često su i rešenje u prevazilaženju problema takozvane "poslednje milje", koriste se za dovođenje interneta u ruralna područja i međusobno povezivanje mobilnih baznih stanica.

Glavni nedostatak FSO sistema je slabljenje optičkog signala u slobodnom prostoru usled prisustva nepovoljnih atmosferskih uslova (kiša, sneg, uslovi slabe vidljivosti) [4]. Apsorpcija i rasejanje svetlosti u magli, oblacima, prašini i dimu značajno slabe laserski snop i ograničavaju performanse sistema i dostupnost optičke veze [1, 2, 4, 5]. Velike varijacije slabljenja mogu čak prouzrokovati relativno duge periode prekida optičke veze.

Performanse optičkog sistema koji radi u uslovima turbulencije mogu se definisati ukoliko je poznat matematički model funkcije gustine verovatnoće (pdf - probability density function) zračenja optičkog signala (scintilacije) [2], [5]. Prema tome, prvi korak u proučavanja prenosa optičkog talasa kroz slobodni prostor je identifikacija pdf-a koji opisuje promenu intenziteta signala u svim uslovima turbulencije. Tokom godina, u literaturi se pojavljuje veći broj različitih teorijskih modela [4], [6]-[12], ali još uvek ne postoji jedinstvena raspodela koja opisuje varijacije intenziteta prenosnog signala u svim režimima turbulencije.

Raspodele koje su najčešće korišćene u literaturi su lognormalna (LN - lognormal) [4], inverzna Gausova kao manje složena raspodela u poređenju sa LN raspodelom [6], takozvana Beckmanova raspodela [7], Gama-Gama (GG - Gamma-Gamma) [8], eksponencijalno-Vejbulova (EW - exponentiated Weibull) [9], [10] i Malaga (M) raspodela [11], [12]. LN model dobro opisuje optički signal u uslovima slabe turbulencije, dok je GG model odličan u uslovima srednjih do visokih turbulencija, a takođe dobro aproksimira promene signala u uslovima slabe turbulencije. Bez obzira na to što je Beckmanova raspodela pokazala dobro poklapanje sa eksperimentalnim rezultatima, zbog kompleksne matematičke forme nije bila povoljna za širu primenu u analitičkoj analizi FSO sistema. U novije vreme, kao alternativa LN i GG raspodelama, predloženi su EW i M modeli raspodele. M model je predložen kao generalni statistički model koji dobro opisuje i sferne i ravanske talase u različitim režimima turbulencije.

Neke od strategija koje se mogu koristiti za prevazilaženje scintilacionih efekata signala uključuju povećanje prenosne optičke snage, upotrebu različitih talasnih dužina, korišćenje više predajnika/prijemnika. U svakom slučaju, veoma je važno tačno opisati uslove turbulencije u FSO kanalu, kako bi se ublažili negativni efekti i poboljšale performanse posmatranog optičkog sistema. U radovima [9] i [10], EW je predložen kao pogodan model za opis realnog kanala turbulencije između krovova dve zgrade duž naseljenog terena srednje gustine u Barseloni. Prikazana je eksperimentalna i uporedna analiza sa ostalim modelima turbulencije, zajedno sa rezultatima simulacije. U radu [13], zabeležena je eksperimentalna validacija FSO kanala u priobalnom okruženju modelovanjem koeficijenta slabljenja u kanalu u zavisnosti od temperature i relativne vlažnosti vazduha, i temperature kondenzacije. Određivanje slabljenja FSO veze u uslovima tropske kiše dato je u radu [14], a eksperimentalna verifikacija modela magle u turbulentnom kanalu u laboratorijskim uslovima u radu [15].

U ovom radu prikazani su rezultati dobijeni u laboratorijski kontrolisanim uslovima u zatvorenom prostoru za

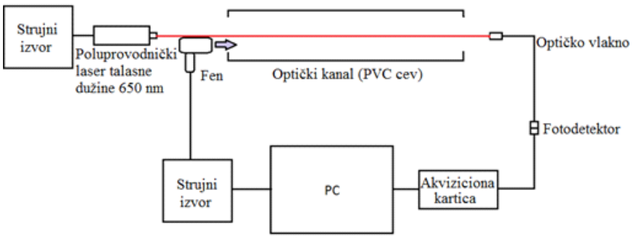
Dejan Milić – Elektronski fakultet, Univerzitet u Nišu, Aleksandra Medvedeva 14, 18115 Niš, Srbija, (e-mail: dejan.milic@ elfak.ni.ac.rs).
Jelena Anastasov – Elektronski fakultet, Univerzitet u Nišu, Aleksandra Medvedeva 14, 18115 Niš, Srbija, (e-mail: jelena.anastasov@ elfak.ni.ac.rs).
Daniela Milović – Elektronski fakultet, Univerzitet u Nišu, Aleksandra Medvedeva 14, 18115 Niš, Srbija, (e-mail: daniela.milovic@ elfak.ni.ac.rs).
Nenad Milošević – Elektronski fakultet, Univerzitet u Nišu, Aleksandra Medvedeva 14, 18115 Niš, Srbija, (e-mail: nenad.milosevic@elfak.ni.ac.rs).

modelovanje slabljenja signala u FSO kanalu. Detaljan opis eksperimenta i deo dobijenih rezultata prikazan je i u radu [16]. U eksperimentalne svrhe, korišćena je kratka turbulentna komora, dužine 3 m, napravljena od PVC cevi. Modulisani predajni signal je emitovan duž komore u kojoj su prisutni efekti turbulencije. Primljeni signal je demodulisan. Izmereni podaci su obrađeni i dobijeni histogram tj. pdf inteziteta primljenog optičkog signala upoređivan sa modelima koji su već predloženi u literaturi. Dobijeni rezultati pokazuju najbolje poklapanje sa EW raspodelom.

II. Eksperimentalna postavka

Šematski prikaz izvedenog eksperimenta dat je na Sl. 1. Postavka se sastoji od tri glavna dela:

- izvor svetlosti
- komora za turbulenciju
- fotodetektor sa sistemom za akviziciju podataka



Sl. 1. Šematski prikaz eksperimenta

Na Sl. 2. prikazana je fotografija eksperimenta u Laboratoriji za telekomunikacione sisteme Univerziteta u Nišu, na Elektronskom fakultetu.

Kao izvor svetlosti korišćen je poluprovodnički laser. Talasna dužina emitovane svetlosti je 650 nm, što je u crvenom delu vidljivog spektra, pa je snop vidljiv i zbog toga se njegov položaj može lako prilagoditi. Korišćen je izvor konstantne struje za pobudu laserskih dioda. Injekciona struja je prilagođena u opsegu od nula do 50 mA, što odgovara emitovanoj optičkoj snazi između nula i 10 mW. Strujni prag laserskog zračenja je detektovan na 10,5 mA, na sobnoj temperaturi. Optička snaga izmerena je pomoću termopilot detektora koji je dostupan u laboratoriji, a merenja su potvrdila dobru stabilnost emitovane snage nakon izvesnog vremena potrebnog da se temperatura stabilizuje.

Turbulentna komora, unutrašnjeg prečnika 20 cm izgrađena je korišćenjem više međusobno povezanih PVC cevi, ukupne dužine 3 m. Duž komore raspoređena su četiri ventilatora (fena), pri čemu je svaki ventilator usmeren na osu komore, tj. na smer prostiranja laserskog snopa. Ventilatori se kontrolišu pojedinačno u smislu nivoa napona napajanja, što rezultira kontrolom brzine kojom se okreću, a time i kontrolom brzine protoka vazduha koji proizvode. Pored toga, temperaturni gradijent je dodat pomoću fena za kosu koji duva vrući vazduh na jednom kraju komore. Pored toplotnog gradijenta, fen dodaje snagu turbulencije svojim dodatnim protokom vazduha. Promenom ugla pod kojim je fen usmeren u odnosu na putanju širenja laserskog zraka mogu se postići različiti uslovi turbulencije, zahvaljujući različitim gradijentima temperature i uslovima protoka vazduha.

Sl. 2. Fotografija eksperimenta iz laboratorije

Laserski zrak pada na čeonu površinu konektora multimodnog optičkog vlakna, a zatim se vodi kratkim vlaknom do prijemnika sa PIN fotodiodom. Optički prijemnik je dizajniran i izrađen u laboratoriji i povezan sa sistemom za akviziciju. Sistem akvizicije zasnovan je na Arduino platformi i koristi svoj interni ADC (analogno-digitalni) pretvarač kako bi uzorkovao analogni signal iz prijemnika. Nakon svake konverzije, mikrokontroler prenosi podatke putem USB veze na host računar gde se podaci čuvaju za dalju obradu.

III. Interpretacija eksperimentalnih podataka

Na početku eksperimenta utvrđeni su referentni nivoi signala testiranjem kombinacije predajnik/prijemnik bez ikakvih turbulencija u prenosnom kanalu. Prijemnik je tipa visoke impedanse. ADC rezolucija je 4,9 mV. Prijemnik je ograničen nivoom šuma predpojačavača koji iznosi $40\text{nV}/\sqrt{\text{Hz}}$, a za propusni opseg koji se razmatra može se prevesti u ulaznu struju šuma od 16 nA. Zbog toga je nivo šuma prijemnika ispod granice ADC rezolucije i prijemnik je u stanju da precizno detektuje konstantnu optičku snagu.

Pojam turbulencije često se karakteriše normalizovanom varijansom fluktuacija zračenja [5] ili indeksom scintilacije, kao

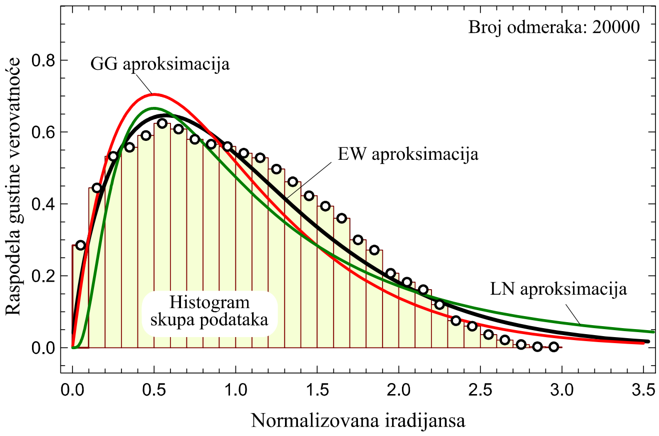
$$\sigma_I^2 = \frac{\mathrm{E}[I^2]}{\left(\mathrm{E}[I]\right)^2} - 1, \tag{4}$$

gde je I nivo zračenja na prijemnoj strani, a E[.] je operator matematičkog očekivanja [17]. Zračenje je direktno proporcionalno izmerenoj fotostruji, a samim tim i proporcionalno procesa na izlazu sistema akvizicije, X. Momenti stohastičkog procesa X se estimitraju na osnovu skupa odmeraka $\{x_i, 1 \le i \le N\}$ i mogu se definisati na sledeći način

$$\mathrm{E}[X^k] = \frac{1}{N}\sum_{i=1}^{N} x_i^k, \tag{5}$$

gde je indeks scintilacije $\sigma_I^2 = \mathrm{E}[X^2]/\left(\mathrm{E}[X]\right)^2 - 1$, sve dok je proces ergodičan.

Kako se ventilatori uključuju u raznim kombinacijama, uspostavljaju se različiti uslovi turbulencije, a nivoi primljenog signala pokazuju manje ili veće varijacije. Nakon prikupljanja velikog broja uzoraka, podaci su obrađeni u cilju definisanja statističkih parametara signala. Da bismo procenili pdf koji opisuje detektovane nivoe signala, iz svake serije uzoraka konstruisan je histogram, a zatim normalizovana skala tako da se ispuniti uslov normalizacije verovatnoće (Sl. 3). Ovako dobijeni pdf je dalje upoređivan sa postojećim modelima u literaturi. Na Sl. 3, prikazane su i krive funkcije gustine verovatnoće LN, GG i EW modela.



Sl. 3. Histogram prikupljenih podataka, normalizovan za procenu pdf-a Indeks scintilacije je $\sigma_I^2 = 0.3614$

IV. Pregled modela turbulencije u literaturi

Idealan teorijski model koji opisuje promene nivoa prenosnog signala bio bi onaj koji je validan u svim režimima turbulencije, za bilo koju veličinu prijemne aperture. Potreban je zatvoreni matematički oblik modela, a da njegovi parametri budu direktno povezani sa fizičkim parametrima atmosferskih uslova. Nažalost, takav savršeni model još uvek nije poznat, ako postoji. Neki od najčešće korišćenih modela u literaturi definisani su u ovom odeljku i upoređeni sa dobijenim eksperimentalnim rezultatima u sledećem.

LN model je najčešće korišćeni model u uslovima slabe turbulencije [1], [4]. Dobijen je na osnovu prvog reda Ritove aproksimacije i pdf fluktuacija inteziteta signala ima sledeći oblik

$$p_I(x) = \frac{1}{x\sigma_i\sqrt{2\pi}}\exp\Big[-\frac{(\log x + \sigma_i^2/2)^2}{2\sigma_i^2}\Big], \quad x > 0 \qquad (6)$$

gde je σ_i^2 varijansa normalizovanog log-zračenja i $E[\log I] = -\sigma_i^2/2$. Indeks scintilacije koji odgovara LN modelu turbulencije se izračunava kao $\sigma_{I_{LN}}{}^2 = -1 + \exp(\sigma_i^2)$.

Nakagami-*m* model koji je poznat u karakterizaciji radio prenosa može se primeniti i u opisu prenosa optičkog talasa [7] kao Backman/modifikovani Rician model turbulencije. Funkcija gustine verovatnoće normalizovanog intenziteta signala, *I*, u tom slučaju, ima sledeći oblik

$$p_I(x) = \frac{1}{b}\exp\Big(-\frac{1+x}{b}\Big)I_0\Big(2\frac{\sqrt{x}}{b}\Big), \quad b, x > 0, \qquad (7)$$

gde je $I_0(.)$ modifikovana Beselova funkcija prve vrste [17]. Model je validan za indekse scintilacije ispod 2. Teorijski indeks scintilacije je $\sigma_{I_{MR}}{}^2 = -1 + \frac{2}{(1+1/b)^2}L_2(-1/b)$, gde je $L_2(x)$ Lagerov polinom drugog reda $L_2(x) = x^2/2 - 2x + 1$.

GG model turbulencije [2], [8] predstavlja stohastički model baziran na teoriji scintilacije. Ovaj model podrazumeva fluktuacije inteziteta signala kao proces modulacije koji nastaje zbog small-scale i large-scale vrtloga turbulencije. I small-scale i large scale efekti se opisuju Gama raspodelom, odakle i potiče ime ovog modela turbulencije. Konačni pdf normalizovane iradijanse, *I*, ima sledeći oblik

$$p_I(x) = \frac{2(\alpha\beta)^{(\alpha+\beta)/2}}{\Gamma(\alpha)\Gamma(\beta)}x^{(\alpha+\beta)/2-1}K_{\alpha-\beta}(2\sqrt{\alpha\beta x}), \quad \alpha, \beta, x > 0 \qquad (8)$$

gde je $\Gamma(.)$ Gama funkcija i $K_\nu(.)$ Beselova funkcija drugog reda [17]. Parametri α i β su parametri broja efektivnih large-scale i small-scale sketerera, respektivno, koji su u direktnoj vezi sa atmosferskim parametrima C_n^2 i l_0. Indeks scintilacije se može izračunati korišćenjem $\sigma_{I_{GG}}{}^2 = \Big(1+\frac{1}{\alpha}\Big)\Big(1+\frac{1}{\beta}\Big) - 1$.

M statistički model [11], [12] je predložen kao generalizovani model koji se može svesti na skoro sve prethodno pomenute modele u literaturi. Formulacija ovog modela je u vidu sledećeg pdf-a iradijanse *I*

$$p_I(x) = A^{(G)}\sum_{k=1}^{\infty} a_k^{(G)} x^{\frac{\alpha+k}{2}-1} K_{\alpha-k}\Big(2\sqrt{\frac{\alpha x}{\gamma}}\Big), \quad \alpha, \theta, \gamma, x > 0 \qquad (9)$$

gde je $A^{(G)} = \frac{2}{\gamma\Gamma(\alpha)}\Big(\frac{\alpha}{\gamma}\Big)^{\frac{\alpha}{2}}\Big(\frac{\gamma\theta}{\gamma\beta+1}\Big)^{\theta}$ i $a_k^{(G)} = \frac{(\theta)_{k-1}(\alpha\gamma)^{\frac{k}{2}}}{\big[(k-1)!\big]^2\gamma^{k-1}(\gamma\theta+1)^{k-1}}$, a $(t)_i$ Pohamerov simbol i θ parameter Nakagami-*m* raspodele.

EW model turbulencije [9], [10] je pokazao odlično poklapanje sa rezultatima simulacije i eksperimentalnim rezultatima u uslovima slabih do srednjih turbulencija u FSO kanalu u svim uslovima usrednjavanja aperture. Još jedna karakteristika ovog modela je jednostavan zatvoreni analitički oblik pdf-a, koji se može prikazati kao [9]

$$p_I(x) = \alpha\beta x^{\beta-1}\exp(-x^\beta)\big[1-\exp(-x^\beta)\big]^{\alpha-1}, \quad \alpha, \beta, x > 0, \qquad (10)$$

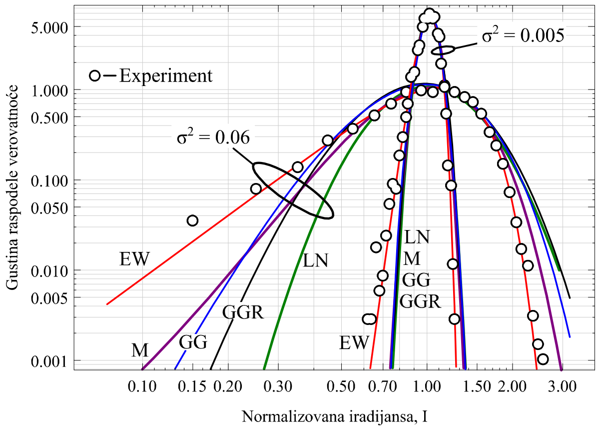
Indeks scintilacije ima oblik $\sigma_{I_{EW}}{}^2 = \frac{\Gamma(1+2/\beta)g_2(\alpha,\beta)}{\alpha\big(\Gamma(1+1/\beta)g_1(\alpha,\beta)\big)^2} - 1$, gde je $g_n(\alpha,\beta) = \Gamma(\alpha)\sum_{j=0}^{\infty}\frac{(-1)^j}{j!(j+1)^{1+n/\beta}\Gamma(\alpha-j)}$.

V. Eksperimentalni rezultati i diskusija

Promenom uglova i brzina ventilatora dobijeni su uslovi u kanalu koji rezultiraju vrednostima indeksa scintilacije u rasponu od 0.0053 do 0.3614, koji su stoga klasifikovani kao uslovi slabe turbulencije. Uslovi jake turbulencije zahtevaju duže prenosne deonice i/ili jače ventilatore. Koristeći prikupljene podatke, vršili smo upoređivnje dobijenog modela sa modelima navedenih u prethodnom odeljku. Prikazane krive su fitovane na osnovu kriterijuma najmanje srednje kvadratne vrednosti, pozivanjem Levenberg-Markuardt algoritma [18].

Na Sl. 4 i 5 prikazane su funkcije gustine verovatnoće normalizovane iradijanse *I* za različite indekse scintilacije.

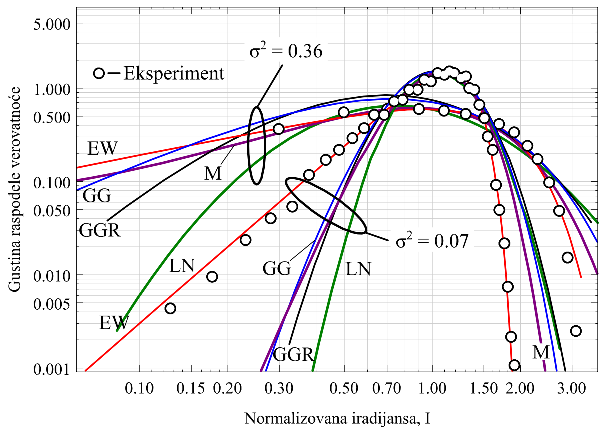
Suprotno očekivanjima, LN model je pokazao najslabije poklapanje sa dobijenim rezultatima osim u uslovima najslabije turbulencije. Ovo je posebno važno kada se razmatraju pdf repovi, jer verovatnoća repa ima veliki značaj u proceni performansi telekomunikacionih sistema. U tom delu grafika, vidno je odstupanje LN modela u odnosu na dobijeni eksperimentalni model.



Sl. 4. Eksperimentalni rezultati dobijeni u režimu slabe turbulencije, upoređeni sa modelima iz literature kada je indeks scintilacije $\sigma_I^2 = 0.06$ i $\sigma_I^2 = 0.005$

Nešto bolje poklapanje se postiže u poređenju sa GG modelom. Razmatrana su dva slučaja: jedan koji odgovara parametrima prema Ritovoj teoriji za ravne talase (GGR), a drugi koji odgovara slobodnom izboru vrednosti parametara (GG). Primećeno je da u najslabijim uslovima turbulencije postoji mala razlika između ova dva slučaja. Sa slika se uočava bolje poklapanje sa eksperimentalnim rezultatima kada indeks scintilacije raste. Drugi razmatrani slučaj GG modela daje bolje rezultate, što dovodi do zaključka da Ritov-ova aproksimacija ravnog talasa nije validna u potpunosti za uslove koji se ispituju. Kada su parametri u GG modelu nezavisni, rezultati pokazuju da su fitovane vrednosti parametra α mnogo veće od vrednosti parametra β. Ovo potvrđuje da je laserski snop koji se koristi samo delimično koherentan i da je dužina koherentnosti relativno mala [19].

Sa slika se može primetiti vidno odstupanje M modela od izmerenih podataka, što ukazuje na veliku kompleksnost ovog modela. Tačnije, postoje značajne numeričke poteškoće pri pokušaju da se podaci o slaboj turbulenciji prilagode ovom opštem M modelu. Slični problemi su primećeni i za opšti GG model u uslovima slabe turbulencije.



Sl. 5. Eksperimentalni rezultati dobijeni u režimu slabe turbulencije i upoređeni sa modelima iz literature kada je indeks scintilacije $\sigma_I^2 = 0.36$ i $\sigma_I^2 = 0.07$

Rezultati pokazuju da se EW raspodela najbolje poklapa sa izmerenim eksperimentalnim rezultatima. U analizi su parametri ove raspodele birani nezavisno s obzirom na to postoji ograničenja za definisanje parametara u različitim atmosferskim uslovima. Za niže vrednosti normalizovane iradijanse, EW grafik se dobro slaže sa dobijenim grafikom, dok je to slaganje za veće vrednosti *I* skoro pa odlično.

VI. Zaključak

U ovom radu prikazan je deo originalnih rezultata dobijenih u laboratorijskim uslovima u cilju karakterizacije pdf-a iradijanse signala u FSO kanalu sa turbulencijama. Analiza je pokazala da je EW model turbulencije adekvatan matematički model za izmerene rezultate pod datim ograničenjima. Uslovi turbulencije u opisanom laboratorijskom okruženju zatvorenog prostora samo su približno reprezentativni za atmosferske uslove u realnim komunikacionim FSO sistemima. Zbog toga bi dalja istraživanja u ovoj oblasti trebalo usmeriti ka izvođenju eksperimenata u uslovima turbulencije na otvorenom i pri prenosu na dužim deonicama FSO kanala.

Zahvalnica

Ovaj rad je podržan od strane Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije.

Literatura

[1] V.I. Tatarskii: Wave Propagation in Turbulent Atmosphere [in Russian], Moscow, SSSR: Nauka, 1967.
[2] L. Andrews, R. Philips: Laser Beam Propagation through Random Media, 2nd ed., Bellingham, Washington, USA: Spie Press, 2005.
[3] M. Toyoshima, H. Takenaka, Y. Takayama: Atmospheric turbulence-induced fading channel model for space-to-ground laser communications links, Optics Express, Vol. 19, No. 17, 2011, pp. 15965 – 15975.
[4] X. Zhu, J.M. Kahn: Free-Space Optical Communication through Atmospheric Turbulence Channels, IEEE Transactions on Communications, Vol. 50, No. 8, Aug. 2002, pp. 1293 – 1300.

[5] L.C. Andrews, R.L. Philips, C.Y. Hopen: Laser Beam Scintillation with Applications, Bellingham, Washington, USA: Spie Press, 2001.
[6] N.D Chatzidiamantis, H.G. Sandalidis, G.K. Karagiannidis, M. Matthaiou: Inverse Gaussian modeling of turbulence-induced fading in free-space optical systems, Journal of Lightwave Technology, Vol. 29, No. 10, May 2011, pp. 1590 – 1596.
[7] J.H. Churnside, S.F. Clifford: Log-normal Rician probability density function of optical scintillations in the turbulent atmosphere, Journal of the Optical Society of America A, Vol. 4, No. 10, 1987, pp. 1923 – 1930.
[8] M.A. Al-Habash, L.C. Andrews, R.L. Phillips: Mathematical model for the irradiance probability density function of a laser beam propagating through turbulent media, Optical Engineering, Vol. 40, No. 8, Aug. 2001, pp. 1554 – 1562.
[9] R. Barrios, F. Dios: Exponentiated Weibull distribution family under aperture averaging for Gaussian beam waves, Optics Express, Vol. 20, No. 12, 2012, pp. 13055 – 13064.
[10] R. Barrios, F. Dios: Exponentiated weibull model for the irradiance probability density function of a laser beam propagating through atmospheric turbulence, Optics & Laser Technology, Vol.45, Feb. 2013, pp. 13 – 20.
[11] A. J-Navas, J.M. G-Balsells, J.F. Paris, A. P-Notario: A Unifying Statistical Model for Atmospheric Optical Scintillation, Numerical Simulations of Physical and Engineering Processes, InTech, DOI: 10.5772/25097, July 2011.
[12] J.M. G-Balsells, A. J-Navas, J.F. Paris, M. C-Vázquez, A. P-Notario: Novel formulation of the M model through the Generalized-K distribution for atmospheric optical channels, Optics Express, Vol. 23, No. 5, 2015, pp. 6345 – 6358.
[13] W.G Alheadary, K.-H. Park, N. Alfaraj, Y. Guo, E. Stegenburgs, T.K. Ng, B.S. Ooi, M.-S. Alouini: Free-space optical channel characterization and experimental validation in a coastal environment, Optics Express, Vol. 26, No. 6, 2018, pp. 6614 – 6628.
[14] S.A. Al-Gailani, A.B. Mohammad, U.U. Sheikh, R.Q. Shaddad: Determination of rain attenuation parameters for free space optical link in tropical rain, Optik, Vol. 125, No. 4, Feb. 2014, pp. 1575 – 1578.
[15] M. Ijaz, Z. Ghassemlooy, H. Le-Minh, S. Zvanovec, J. Perez, J. Pesek, O. Fiser: Experimental validation of fog models for FSO under laboratory controlled conditions, 24th IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), London, UK, 8-11 Sept. 2013, pp. 19 – 23.
[16] D. N. Milić, D. M. Milović, J. A. Anastasov, "Experimental indoor validation of weak intensity fluctuations over free-space optical communication link", Serbian Journal of Electrical Engineering, Vol. 18, No. 1, February 2021, pp. 15-27.
[17] I.S. Gradshteyn, I.M. Ryzhik: Table of Integrals, Series, and Products. 6th ed., New York: Academic, 2000.
[18] J. More: The Levenberg-Marquardt Algorithm: Implementation and Theory Numerical Analysis, Springer, Berlin Heidelberg, 1977.
[19] O. Korotkova, L.C. Andrews, R.L. Phillips: Model for a partially coherent Gaussian beam in atmospheric turbulence with application in Lasercom, Optical Engineering, Vol. 43, No. 2, Feb. 2004, pp. 330 – 341.

ABSTRACT

In this paper, experimental results obtained in laboratory - controlled conditions for the characterization of turbulence over indoor free space optical (FSO) link are presented. The collected data is used for statistical characterization of the probability density function (pdf) of fluctuations of transmitted optical signal. There are several empirical models in the literature for different turbulence regimes. The obtained pdf is compared with selected models from the literature in order to define the statistical properties of the real turbulent channel, which is important in determining the performance of the FSO system. The analysis has shown that the results obtained under the given conditions coincide well with the exponential - Weibull model of turbulence.

Experimental Turbulence Characterization over Free-space Optical Communication Link

Dejan Milić, Daniela Milović, Jelena Anastasov, Nenad Milošević

Pregled postojećih DPD modela sa ograničenom širinom propusnog opsega

Tamara Muškatirović-Zekić, *Elektrotehnički fakultet Beograd, Regulatorna agencija za elektronske komunikacije i poštanske usluge,* Milan Čabarkapa, *Elektrotehnički fakultet Beograd,* Nataša Nešković, *Elektrotehnički fakultet Beograd,* Đurađ Budimir, *University of Westminster, London*

***Apstrakt*—Novi zahtevi za sve bržim i pouzdanijim servisima i uslugama putem mobilnih mreža dovode do potrebe za razvijanjem sofisticiranijih tehnika digitalne predistorzije (DPD) pojačavača (*power amplifier* - PA). Stoga je u ovom radu dat pregled postojećih DPD modela sa ograničenom širinom propusnog opsega (*band-limited* DPD). Ukratko su opisana predložena rešenja za *band-limited* DPD i data je njihova komparativna analiza. Takođe je izvedena i analiza kompleksnosti predloženih modela.**

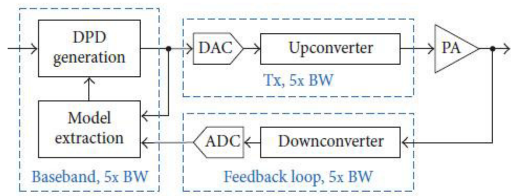
***Ključne reči*— *band-limited* DPD, PA, LDMOS Doherty PA, LTE, PAPR, MIMO, mMIMO.**

I. UVOD

SVE je veći broj pametnih uređaja koji se putem mobilnih mreža nove generacije (4G i 5G) povezuju na Internet, pružajući korisnicima nove servise i istovremeno povećavajući potrebu za prenosom velike količine podataka, velikim brzinama. Očekuje se da će 5G pristupne mreže zahtevati signale širokog propusnog opsega, reda stotinak MHz, pa sve do nekoliko GHz [1]. Usled porasta propusnog opsega (*bandwidth*), povećavaju se i zahtevi koji se stavljaju pred pojačavač (PA), koji predstavlja najveći potrošač u predajniku, i DPD (*Digital PreDistortion*) sistem koji se koristi u cilju održavanja linearnosti i poboljšanja efikasnosti rada PA.

Kod konvencionalnih DPD sistema, potrebna širina propusnog opsega je minimum 5 puta veća od širine signala, kako bi se kompenzovala nelinearna distorzija usled širenja spektra. Razlog je taj što većina *predistorted* signala obuhvata pored osnovne komponente i intermodulacione proizvode trećeg i petog reda [2]. Tako se za signal širine 500 MHz dobija zahtevana širina propusnog opsega od 2500 MHz, odnosno brzina odabiranja DAC i ADC od 3200 Msps, za vrednost roll-of faktora od 0.28 [3]. Konvertori sa ovako velikim brzinama odabiranja su izuzetno skupo i energetski neisplativo rešenje. Cilj je realizovati širokopojasni DPD sa A/D konvertorom male brzine *(low-speed ADC*).

Na slici 1 je prikazana blok šema konvencionalnog DPD sistema.



Sl. 1. Konvencionalni DPD sistem [3].

II. DPD SISTEMI SA OGRANIČENOM ŠIRINOM PROPUSNOG OPSEGA

Većina istraživača se fokusirala na smanjenje širine propusnog opsega u povratnoj grani.

U [4] i [5] je predložen *band-limited* DPD model koji rešava problem širine propusnog opsega u Tx i povratnoj grani. Kod ovog modela je *band-limiting* funkcija ubačena u *Volterra* operator u DPD modelu, kako bi se kontrolisala širina propusnog opsega izlaznog signala. *Band-limited Volterra* operator *p*-tog reda se može predstaviti sa:

$$T_p\left(x(n)\right) = D_p\left(x(n)\right) \otimes f(n) \qquad (1)$$

gde je: D_p- *Volterra* operator *p*-tog reda, *f* - *band-limiting* funkcija, a $\otimes$ predstavlja operaciju konvolucije. Za *band-limiting* funkciju je korišćen 40 MHz *low-pass complex equiripple* FIR filter. Predložen DPD model je eksperimentalno verifikovan u više testova na *high-power* LDMOS Doherty PA na 2.14 GHz. Izvršeno je poređenje konvencionalnog DPD modela sa potrebnom širinom propusnog opsega, konvencionalnog DPD modela sa ograničenom širinom propusnog opsega i predloženog DPD modela. Korišćen je WCDMA signal širine 20 MHz sa 4 nosioca i PAPR=6.5 dB. Zatim je izvršena eksperimentalna provera uticaja promene propusnog opsega sistema na performanse linearizacije, pri čemu je korišćen UMTS signal širine 60 MHz sa 12 nosioca i PAPR=6.5 dB. I na kraju je eksperimentalno proveren predložen DPD model kod sistema sa signalima šireg propusnog opsega i sa različitim kofiguracijama signala. Teorijska analiza i eksperimentalni rezultati u [5] su pokazali da ova tehnika značajno poboljšava performanse sistema i smanjuje troškove implementacije

Tamara Muškatirović-Zekić – Elektrotehnički fakultet u Beogradu, Bulevar kralja Aleksandra 73, 11120 Beograd, Srbija; Regulatorna agencija za elektronske komunikacije i poštanske usluge, Palmotićeva 2, 11000 Beograd, Srbija (e-mail: tamara.muskatirovic@ratel.rs)

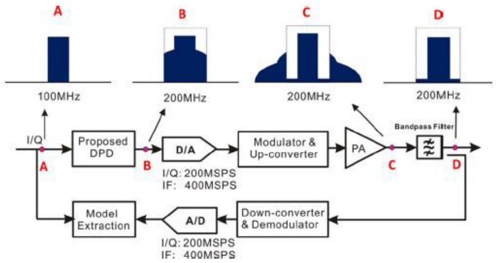
Milan Čabarkapa – Elektrotehnički fakultet u Beogradu, Bulevar kralja Aleksandra 73, 11120 Beograd, Srbija (e-mail: milan.cabarkapa@gmail.com)

Nataša Nešković – Elektrotehnički fakultet u Beogradu, Bulevar kralja Aleksandra 73, 11120 Beograd, Srbija (e-mail: natasha@etf.rs)

Đurađ Budimir – Wireless Communications Research Group, University of Westminster, London W1W 6UW, U.K. (e-mail: d.budimir@wmin.ac.uk).

DPD, i što je još važnije, eliminiše ograničenja u pogledu propusnog opsega sistema koja postoje kod konvencionalnih DPD tehnika.

Na slici 2 je prikazana blok šema *band-limited* DPD sistema.



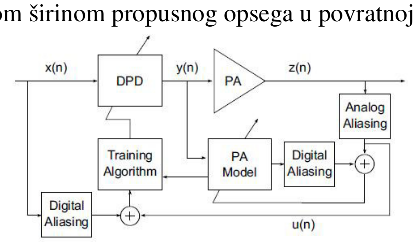
Sl. 2. *Band-limited* DPD sistem [5].

U [6] je predložen LS (*least squares*) model za procenu parametara DPD modela sa ograničenom širinom propusnog opsega, kako bi se poboljšale performanse *band-limited* DPD. Prilagođenje je izvršeno postavljanjem frekvencijskog ograničenja na širinu propusnog opsega konvencionalnog LS modela ekstrakcije. Predloženi model je eksperimentalno verifikovan na širokopojasnom *Doherty* PA izlazne snage 37 dBm, koji je dizajniran za 3G/4G sisteme. Korišćen je pojednostavljen DDR (*Dynamic Deviation Reduction*) *Volterra* model drugog reda i UMTS signal širine 40 MHz sa 8 nosioca i PAPR=6.5 dB. Primenom predloženog modela postignut je ACLR (*Adjacent Channel Leakage Ratio*) ispod -55 dBc.

Određivanje koeficijenata *band-limited* DPD modela u [7] je formulisano kao uopšten LS (GLS - *Generalized Least Squares*) problem posmatrano u frekvencijskom domenu. Ovom metodom je operacija konvolucije u vremenskom domenu zamenjena efikasnom operacijom FFT, čime je značajno smanjena računska kompleksnost. Eksperimentalni rezultati, sa WCDMA signalom širine 20 MHz i LTE signalom širine 40 MHz sa 2 nosioca, su pokazali da su performanse sa FDD-BL-DPD (*Frequency Domain Data Based Band-Limited DPD*) podjednako dobre kao sa TDD-BL-DPD (*Time Domain Data Based Band-Limited DPD*).

U [8] je predložen metod određivanja DPD parametara metodom direktnog učenja (*Direct Learning Algorithm* - DLA) sa ograničenom širinom propusnog opsega u povratnoj grani. Eksperimentalna merenja su pokazala da se predloženim algoritmom mogu postići slične performanse kao kod konvencionalnih DPD sistema, uz značajno smanjenje širine propusnog opsega u povratnoj grani.

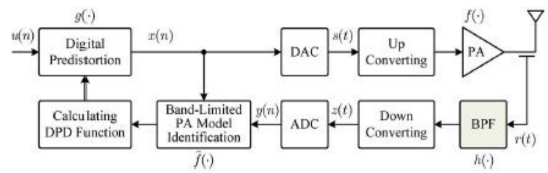
Na slici 3 je prikazana blok šema DLA DPD sistema sa ograničenom širinom propusnog opsega u povratnoj grani.



Sl. 3. DLA sa ograničenom širinom propusnog opsega u povratnoj grani [8].

U [9] je predstavljen DPD model sa ograničenom širinom propusnog opsega u povratnoj grani, kod koga se za određivanje parametara PA koristi *memory polinomial* (MP) model. Dobijeni parametri PA su zatim korišćeni u algoritmu direktnog učenja (DLA) kako bi se dobila DPD funkcija koja se zatim koristi za linearizaciju PA. Model je eksperimentalno potvrđen na LTE-A signalu širine 100 MHz sa 5 nosioca. Eksperimentalni rezultati su potvrdili poboljšanje ACLR od 22 dB, čak i kada se opseg u povratnoj grani ograniči na 100 MHz, a brzina odabiranja ADC značajno smanji na 368.64 Msps, čime se značajno smanjuju teškoće u projektovanju sistema kao i troškovi implementacije.

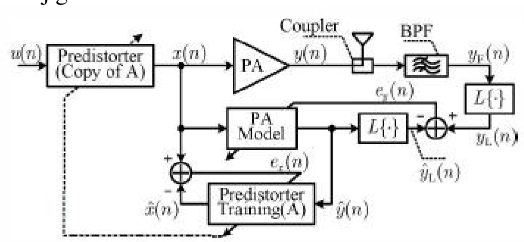
Na slici 4 je prikazana blok šema MP DLA DPD sistema sa ograničenom širinom propusnog opsega u povratnoj grani.



Sl. 4. MP DLA DPD sa ograničenom širinom propusnog opsega u povratnoj grani [9].

Kod DPD modela sa ograničenom širinom propusnog opsega u povratnoj grani u [10] je za određivanje parametara PA korišćen PTNTB *(Parallel Twin Nonlinear Two-Box)* model i algoritam indirektnog učenja (ILA). Eksperimentalni rezultati, na LTE-A signalu širine 100 MHz sa 5 nosioca, pokazali su da predložen model poboljšava ACLR performanse za 23 dB, čak i kada se širina propusnog opsega u povratnoj grani smanji sa 500 MHz na 140 MHz.

Na slici 5 je prikazana blok šema PTNTB ILA DPD sistema sa ograničenom širinom propusnog opsega u povratnoj grani.

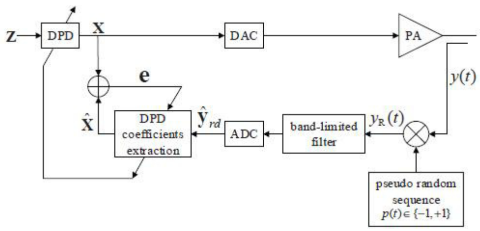


Sl. 5. PTNTB ILA DPD sa ograničenom širinom propusnog opsega u povratnoj grani [10].

Kako bi omogućili korišćenje sporih (*low-speed*) ADC konvertora, autori u [11] su predložili korišćenje spektralne ekstrapolacije signala sa ograničenom širinom propusnog opsega u povratnoj grani. Za optimalno određivanje DPD parametara korišćenog MP modela predložen je *dumped Gauss Newton* algoritam. Eksperimentalno je pokazano da širina propusnog opsega u povratnoj grani može da bude i manja od širine signala, a da se i dalje dobiju dobre performanse.

RDRS (*random demodulation based reduced sampling rate*) metod za modelovanje i linearizaciju PA, predložen u [12], značajno redukuje brzinu odabiranja u odnosu na metod spektralne ekstrapolacije, pri čemu i dalje zadržava dobre performanse. Ovaj metod koristi tehniku slučajne demodulacije u povratnoj grani sa ograničenom širinom propusnog opsega. Značajna karakteristika ove metode je ta što množenje signala pseudoslučajnom sekvencom u vremenskom domenu, širi signal preko celog spektra, pa svaka tačka na frekvencijskoj osi sadrži sve informacije o signalu. Ovo implicira da je za vraćanje signala dovoljno koristiti veoma mali deo spektra.

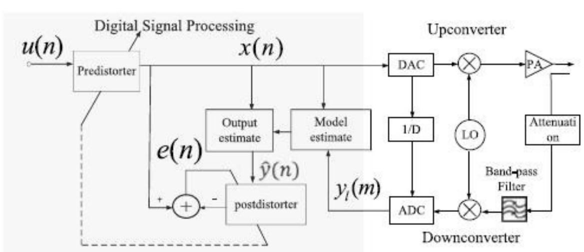
Na slici 6 je prikazana blok šema DPD sistema zasnovanog na slučajnoj demodulaciji (RDRS).



Sl. 6. RDRS DPD model [12].

DPD tehnika sa ultra malom brzinom odabiranja je predložena u [13], i sastoji se iz dva koraka. U prvom koraku se pomoću *high-rate* ulaznog signala i *low-rate* izlaznog signala iz *low-speed ADC* u povratnoj grani određuju koeficijenti modela. U drugom koraku se dobijeni model koristi za procenu *high-rate* izlaznog signala PA. Zatim se na osnovu *high-rate* ulaznog signala i procenjenog *high-rate* izlaznog signala određuju koeficijenti inverznog modela, koji se koristi za kompenzaciju nelinearnosti PA. Ovom tehnikom su značajno smanjeni zahtevi za brzinom ADC u povratnoj grani. Eksperimentalni rezultati na LTE signalu širine 40 MHz su pokazali da se predloženom tehnikom može postići smanjenje brzine ADC konvertora do čak 2.5 Msps, uz postizanje skoro iste linearizacije kao kod prethodno predloženih *band-limited* DPD tehnika, čime su značajno smanjeni zahtevi za širinom propusnog opsega u povratnoj grani.

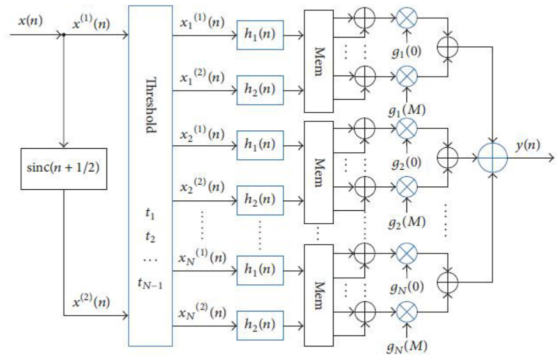
Na slici 7 je prikazana blok šema DPD sistema sa malom brzinom odabiranja.



Sl. 7. DPD sistem sa malom brzinom odabiranja [13].

U cilju daljeg poboljšanja linearizacije PA za *mmWave* frekvencijski opseg, koji se nameće kao jedan od kandidata za 5G mobilne sisteme, autori u [3] su predložili DPD tehniku sa smanjenim zahtevima za širinu propusnog opsega za Tx, povratnu granu, ali i za osnovni opseg (Tx/FB/BB). Predložen DPD model koristi skup linearnih "*piecewise*" segmenata za opis nelinearnih karakteristika PA, zamenjujući operatore višeg reda sa nekoliko operatora nižeg reda. Na taj način se smanjuje zahtev za širinom propusnog opsega u osnovnom opsegu i omogućava se primena modela kod budućih ultra širokopojasnih sistema. Tehnika je eksperimentalno potvrđena na mmWave PA na 41 GHz (*in-housed designed mmWave PA module*). Testovi su vršeni sa LTE signalom širine 80 MHz sa 4 nosioca i PAPR=7.5 dB, i sa LTE signalom širine 320 MHz sa 4 nosioca i PAPR=7.5 dB.

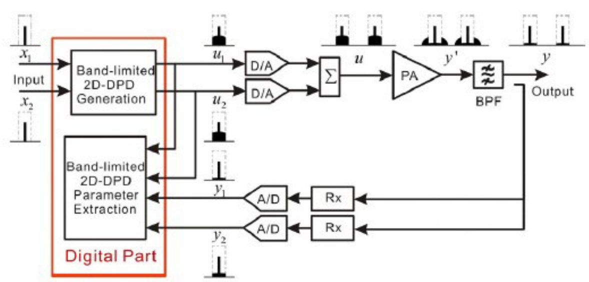
Na slici 8 je prikazana blok šema DPD sistema za ultra-širokopojasne *mmWave* PA.



Sl. 8. DPD sistem za ultra-širokopojasne *mmWave* PA [3].

Za linearizaciju širokopojasnih RF PA u *concurrent dual-band* predajnicima u [14], predložen je BL 2-D (*band-limited two-dimensional*) DPD. Model je eksperimentalno potvrđen na *high-power* LDMOS Doherty PA na 2.14 GHz, sa izlaznom snagom od 37 dBm. *Dual-band* signal je generisan korišćenjem LTE signala širine 20 MHz i WCDMA signala širine 20 MHz sa 4 nosioca. Zbog ograničenja testnog modela, frekvencijski razmak je postavljen na 80 MHz. Maksimalna širina svakog opsega u povratnoj grani je postavljena na 80 MHz. Eksperimentalni rezultati pokazuju da predložen metod predstavlja moguće rešenje za efikasno smanjenje zahteva za širinom propusnog opsega za širokopojasne *concurrent dual-band* predajnike.

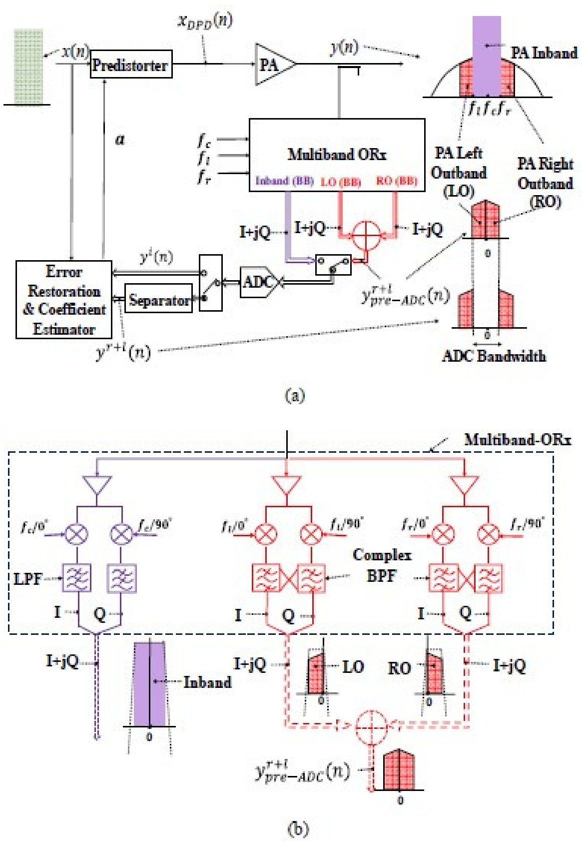
Blok šema predloženog modela prikazana je na slici 9.



Sl. 9. *Band-limited* 2-D DPD sistem [14].

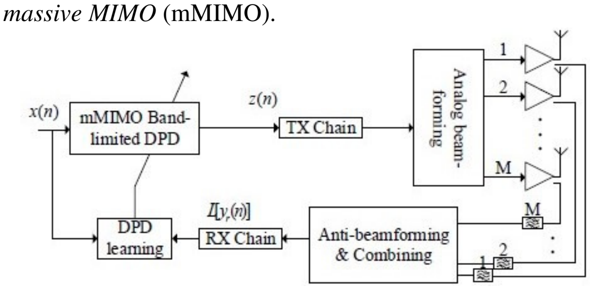
Kod klasičnog BL DPD modela, odbacivanjem signala van opsega PA dolazi do gubitka značajne količine informacija o nelinearnosti PA. Stoga su autori u [15] predložili BL DPD model sa *Band-switching* arhitekturom u povratnoj grani i ažuriranom procedurom linearizacije širokopojasnog 5G *mmWave* PA. U predloženom modelu se uzima u obzir i signal van opsega PA, korišćenjem *band-switching* arhitekture prikazane na slici 10, uz zadržavanje širine propusnog opsega ADC, kako bi se izbeglo povećanje troškova i potrošnje energije. Uveden je mehanizam za prebacivanje kojim se izdvaja ili signal unutar opsega PA ili signal van opsega PA, na osnovu unapred utvrđenog redosleda kroz odgovarajući filter. Za izdvajanje signala unutar opsega PA se koristi LPF (*low-pass* filter), dok se za izdvajanje signala van opsega PA koristi složeni BPF (*band-pass* filter), nakon što se izvrši frekvencijsko pomeranje odgovarajućim lokalnim oscilatorima. Tokom izdvajanja signala van opsega PA, levi i desni PA opsezi se frekvencijski pomeraju tako da se postavljaju jedan pored drugog kao što je prikazano na slici 10 (a) i (b) i kombinuju. Širina dobijenog signala je jednaka širini PA, čime se zadržava ista širina propusnog opsega ADC kao i u klasičnom BL DPD. Zatim se vrši analiza izdvojenih signala kako bi se ažurirali DPD koeficijenti. U prvih nekoliko iteracija, modelovanje DPD se vrši samo pomoću informacija unutar opsega PA. U narednim iteracijama, naizmenično se biraju signali unutar i van opsega PA, i zatim se vrši obrada u bloku *Error Restoration & Coefficient Estimator block*. Eksperimentalno je potvrđeno da se pomoću predloženog modela postiže ACLR od -43.6/-42 dBc i EVM<3% za *Watt-class* PA sa 256QAM modulisanim OFDM signalom širine 800 MHz u opsegu 26 GHz, pri čemu je mehanizam prebacivanja implementiran u MATLAB-u.

Blok šema predloženog DPD modela sa *Band-switching* arhitekturom prikazana je na slici 10.



Sl. 10. BL DPD model sa *Band-switching* arhitekturom [15].

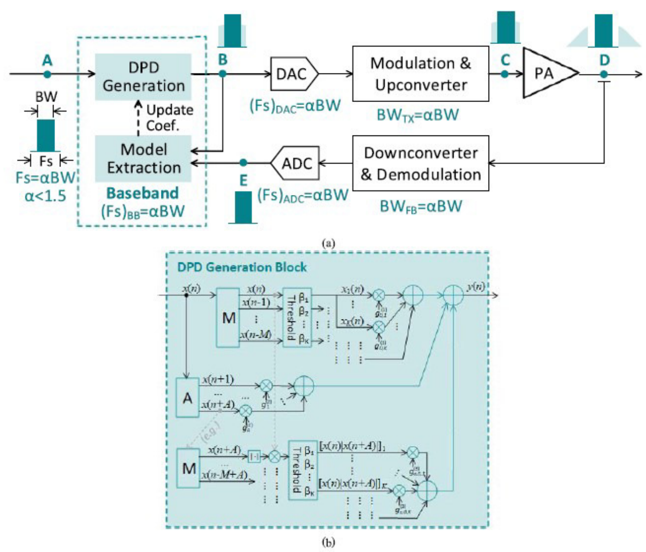
Primena BL DPD modela za hibridne MIMO predajnike analizirana je u [16]. Blok šema predloženog modela dobijena je uvođenjem *band-limiting* funkcije i predstavljena je na slici 11. Model je testiran na *parallel-Wiener* PA modelu, sa ulaznim LTE signalom širine 50 MHz i PAPR=7.6dB. Korišćen je MP DPD model čiji je red nelinearnosti 7, a dubina memorije 3. Za *band-limiting* funkciju je uzet kompleksni BPF propusnog opsega od 60 MHz. Pokazano je da predloženi model postiže performanse uporedive sa postojećim DPD metodama za hibridne MIMO predajnike, ali sa nižim brzinama ADC, čime se može značajno smanjiti cena hardvera uz zadržavanje dobrih performansi pri primeni kod *massive MIMO* (mMIMO).



Sl. 11. mMIMO *band-limited* DPD model [16].

BL DPD model predložen u [17] vrši smanjenje propusnog opsega u osnovnom opsegu primenom linearne dekompozicije. Za razliku od modela predloženog u [13] gde se za interpolaciju koristi *sinc* funkcija koja uključuje dosta množenja, u LD (*Linear-Decomposition*) BL DPD modelu je znatno smanjena računarska kompleksnost uz zadržavanje male brzine odabiranja. Predložen model je eksperimentalno verifikovan na 5G NR test signalu u dva opsega, *sub*-6-GHz (64 QAM signal širine 100 MHz i PAPR=9.1 dB) i *mmWave* (64 QAM signal u 28 GHz, širine 400 MHz i 800 MHz i PAPR= 9dB).

Blok šema predloženog modela data je na slici 12.



Sl. 12. LD BL DPD model [17].

III. POREĐENJE DPD MODELA SA OGRANIČENOM ŠIRINOM PROPUSNOG OPSEGA

Kompleksnost rada pri kojoj izloženi modeli dostižu izložene performanse se dosta razlikuje od modela do modela. Procena računarske kompleksnosti može se izvršiti na osnovu broja FLOP-ova (*floating point operations*). U tabeli 1 je prikazan potreban broj FLOP-ova za određene računarske operacije [18].

TABELA I
POTREBAN BROJ FLOP-OVA ZA RAČUNARSKE OPERACIJE

Operacija	*Broj FLOP-ova*
Konjugovanje	0
Kašnjenje	0
Sabiranje realnih brojeva	1
Množenje realnih brojeva	1
Sabiranje kompleksnih brojeva	2
Množenje kompleksnog i realnog broja	2
$\lvert\cdot\rvert$^2	3
Množenje kompleksnih brojeva	6
Kvadratno korenovanje	7

Procena performansi DPD modela se može izvršiti na osnovu normalizovane srednje kvadratne greške NMSE (*Normalized Mean-Squared Error*), koja se računa po formuli:

$$NMSE = 10\log_{10}\left(\frac{\sum_{n=1}^{K}\lvert y_{meas}(n) - y_{est}(n)\rvert^2}{\sum_{n=1}^{K}\lvert y_{meas}(n)\rvert^2}\right) \quad (2)$$

gde su: y_{meas} i y_{est} - izmereni i procenjeni talasni oblici izlaznog signala, respektivno.

U tabeli 2 je dat uporedni prikaz eksperimentalno izmerenih performansi izloženih modela, kao i izračunata kompleksnost modela, pri čemu je P - red nelinearnosti, M - dubina memorije, K - dužina *band-limiting* funkcije modela, K_a i K_b - red nelinearnosti statičkog nelinearnog modela i MP modela, N - broj "*piecewise*" segmenata, N_s - broj odbiraka u svakoj sekvenci, A - dužina glavnih odbiraka.

IV. ZAKLJUČAK

Na osnovu izloženog, može se zaključiti da iako je sam koncept DPD prilično jednostavan, razvijanje jeftinog i efikasnog DPD modela za potrebe budućih širokopojasnih mobilnih sistema predstavlja veoma izazovan zadatak. Značaj ovog rada je u kompaktnom prikazu i analizi postojećih *band-limited* DPD tehnika. Budući rad ide u pravcu ispitivanja mogućnosti optimizacije i razvijanja efikasnijeg *band-limited* DPD modela.

TABELA II
POREĐENJE PARAMETARA BAND-LIMITED DPD MODELA

Referenca	*Signal*	*Širina signala [MHz]*	*PAPR [dB]*	*Širina propusnog opsega sistema [MHz]*	*ACPR [dBc]*	*NMSE [dB]*	*Kompleksnost modela*
[3]	LTE sa 4 nosioca	80	7.5	144	-47.64/-47.47	-34.63	$2\cdot(4\cdot N\cdot M+10\cdot N+1)$
	CA LTE sa 4 nosioca	320	7.5	576	-42.92/-45.04	-30.13	
[5]	WCDMA sa 4 nosioca	20	6.5	40	-60.44/-60.40	-45.83	$K\cdot\left((M+1)\cdot\frac{P+1}{2}-1\right)+3\cdot K\cdot\left(M\cdot\frac{P-1}{2}-1\right)+3$
	LTE-A sa 5 nosioca	100	7.8	140	-51.02/-52.22	-41.66	
	LTE-A sa 3 nosioca	60	7.7	140	-52.57/-53.44	-41.94	
	LTE-A + UMTS	100	9.2	140	-56.14/-57.04	-43.53	
[8]	WCDMA sa 2 razmaknuta nosioca	40	-	81.92	-59.60/-57.73	-	-
[9]	LTE-A sa 5 nosioca	100	-	100	-48	-	$2M\cdot(13P-7)$
[10]	LTE-A sa 5 nosioca	100	-	140	-48	-	$4\cdot(3K_a+4M(K_b-1)-3)$
[12]	LTE	20	-	20	-52.61/-52.39	-45.77	-
[13]	LTE	40	7.0	20	-46.35/-45.59	-	-
[14]	*Dual-band* LTE i WCDMA sa 4 nosioca	20	-	40	-52.13/-51.54 – (*lower*) 51.15/-50.53 (upp*er*)	-44.46 (*lower*) -43.36 (upp*er*)	-
[15]	*256 QAM OFDM*	800	7.6	800	-43.6/-42	-	-
[16]	*LTE*	50	7.6	60	-65.16/ -64.82	-52.85	-
[17]	*5G NR sub-6GHz*	100	9.1	150	-	-35.26	$2\cdot(4\cdot N\cdot M+10\cdot N+1)$
	5G NR mmWave 28GHz	400	9	600	-	-32.35	
		800	9	1200	-	-29.52	

LITERATURA

[1] 5GPPP, "5G Vision: The 5G Infrastructure Public Private Partnership - The next generation of communication networks and services", February 2015

[2] F. M. Ghannouchi and O. Hammi, “Behavioral modeling and predistortion”, IEEE Microwave Magazine, vol. 10, no. 7, pp. 52–64, December 2009.

[3] C. Yu , Q. Lu, H. Sun, X. Wu and X.-W. Zhu, “ Digital Predistortion of Ultra-Broadband mmWave Power Amplifiers with Limited Tx/Feedback Loop/Baseband Bandwidth”, Wireless Communications and Mobile Computing, vol. 2018, March 2018.

[4] C. Yu, L. Guan and A. Zhu, “Band-Limited Volterra Series-Based Behavioral Modeling of RF Power Amplifiers”, IEEE MTT-S International Microwave Symposium, IMS ’12, pp. 1–3, Canada, June 2012.

[5] C. Yu, L. Guan, E. Zhu and A. Zhu, “Band-Limited Volterra Series-Based Digital Predistortion for Wideband RF Power Amplifiers”, IEEE Transactions on Microwave Theory and Techniques, vol. 60, no. 12, pp. 4198–4208, December 2012.

[6] L. Guan, C. Yu and A. Zhu, “Bandwidth-constrained Least Squares-based Model Extraction for Band-limited Digital Predistortion of RF Power Amplifiers” Conference Paper, in *Workshop INMMiC*, Dublin, Ireland, September 2012.

[7] G. Yang, F. Liu, C. Zhao, Z. Wang, L. Li and H. Wang, “Frequency Domain Data Based Model Extraction for Band-Limited Digital Predistortion of Wideband RF Power Amplifiers” International Journal of RF and Microwave Computer-Aided Engineering, pp. 412–420, August 2013.

[8] L. Ding, F. Mujica and Z. Yang, “Digital Predistortion using Direct Learning with Reduced Bandwidth Feedback,” in Proc. IEEE Int. Microw. Symp., pp. 1–3, Jun. 2013.

[9] Y. Liu, W. Pan, S. Shao and Y. Tang, “A New Digital Predistortion for Wideband Power Amplifiers With Constrained Feedback Bandwidth”, IEEE Microwave And Wireless Components Letters, vol. 23, no. 12, pp. 683–685, December 2013.

[10] Y. Liu, W. Pan, S. Shao, and Y. Tang, “A New Digital Predistortion Using Indirect Learning with Constrained Feedback Bandwidth for Wideband Power Amplifiers”, IEEEMTT-S Int. Microw. Symp. Dig., pp. 1–3, Jun. 2014.

[11] Y. Ma, Y. Yamao, Y. Akaiwa and K. Ishibashi, “Wideband Digital Predistortion Using Spectral Extrapolation of Band-Limited Feedback Signal” IEEE Transactions on Circuits and Systems, vol. 61, no. 7, pp. 2088–2097, 2014.

[12] W. Tao, H. Wang, C. Zhou, G. Li and F. Liu, “A Random Demodulation Based Reduced Sampling Rate Method for Wideband Digital Predistortion,” in Proceedings of the Asia-Pacific Microwave Conference, APMC ’15, China, December 2015.

[13] Z. Wang, W. Chen, G. Su, F. M. Ghannouchi, Z. Feng and Y. Liu, “Low Feedback Sampling Rate Digital Predistortion for Wideband Wireless Transmitters,” IEEE Transactions on Microwave Theory and Techniques, vol. 64, no. 11, pp. 3528–3539, November 2016.

[14] C. Yu, Y. Guo, and A. Zhu, “A Band-limited 2-D Digital Predistorter for Concurrent Dual-band RF Transmitters”, IEEE International Wireless Symposium (IWS), March 2014.

[15] S. Deb, M. Tanio, S. Hori, N. Tawa, Y. Wada, and K. Kunihiro, “Band-limited Digital Predistortion with Band-switching Feedback Architecture for 5G mmWave Power Amplifiers”, 48th European Microwave Conference (EuMC), September 2018.

[16] W. Qiao, G. Li, Y. Zhang, H. Li, and F. Liu, “A Band-limited Digital Predistortion Method for Hybrid MIMO Transmitters”, International Conference on Microwave and Millimeter Wave Technology (ICMMT), May 2019.

[17] C. Yu, Q. Lu, H. Yin, J. Cai, J. Chen, X.-W. Zhu, and W. Hong, “Linear-Decomposition Digital Predistortion of Power Amplifiers for 5G Ultrabroadband Applications”, IEEE Transactions on Microwave Theory and Techniques, vol. 68, no. 7, pp. 2833 - 2844, July 2020.

[18] A.S. Tehrani, C. Haiying, S. Afsardoost, T. Eriksson, M. Isaksson and C. Fager, “A comparative analysis of the complexity/accuracy tradeoff in power amplifier behavioral models,” IEEE Trans. Microw. Theory Tech., vol. 58, no. 6, pp. 1510-1520, June 2010.

ABSTRACT

New demands for faster and more reliable applications and services over mobile networks lead to the need of developing more sophisticated digital pre-distortion (DPD) techniques for power amplifiers (PA). Therefore, this paper provides an overview of existing band-limited DPD models. The proposed solutions for band-limited DPD are briefly described and their comparative analysis is given. An analysis of the complexity of the proposed models was also performed.

An overview of existing band-limited DPD models

Tamara Muškatirović-Zekić, Milan Čabarkapa, Nataša Nešković, Đurađ Budimir

Sistem za detekciju i klasifikaciju niskoletećih besposadnih platformi – dronova (SDKNBP)

Mohammed Mokhtari, Jovan Bajčetić i Boban Sazdić-Jotić

***Apstrakt*—Istraživanje predstavljeno u ovom radu prikazuje mogućnosti planiranog koncepta sistema za detekciju i klasifikaciju niskoletećih besposadnih platformi (dronova) zasnovanom na metodama dubokog učenja. Cilj projekta predstavlja razvoj upotrebljivog sistema koji će u realnom vremenu vršiti detekciju i klasifikaciju dronova na bazi karakteristika njihovih radio emisija. Metode detekcije i klasifikacije primenjene u ovom istraživanju zasnivaju se na konvolucionoj neuronskoj mreži istreniranoj upotrebom formirane baze snimaka radio emisija sa nekoliko vrsta komercijalno dostupnih dronova. U početnoj fazi istraživanja sistem pokazuje verovatnoću detekcije i klasifikacije od 100 % za ukupno četiri nezavisne klase – nema drona, dron 1, dron 2 i dron 3 što predstavlja osnovu za dalji razvoj sistema za detaljniju klasifikaciju.**

***Ključne reči*—Duboko učenje; Detekcija; Klasifikacija; Radio emisija; Dron.**

I. Uvod

JEDNA od najaktuelnijih oblasti tehnološkog i industrijskog razvoja današnjice je oblast razvoja i unapređenja besposadnih platformi. S obzirom na to da se autonomna vozila smatraju zamajcem četvrte industrijske revolucije, posebnu pažnju nauke, tehnologije i industrije privlače leteće besposadne platforme. One će, beskompromisno u skorijoj budućnosti zauzeti vrlo veliki i važan deo ljudske delatnosti primarno u oblasti transporta, ali i u okviru ostalih oblasti delovanja.

Terminologija označavanja niskoletećih besposadnih platformi je pretežno kod nas zastupljena na engleskom jeziku, tako da se u literaturi može naći više različitih terminoloških izraza, kao što su:

- UAS (Unmanned Aircraft System), [1];
- RPAS (Remotely Piloted Aircraft System), [2];
- UAV (Unmanned Aerial Vehicle).

U dokumentu koji koristi francuski direktorat za civilnu avijaciju [2], a taj izraz je široko prihvaćen u svetskoj naučnoj zajednici, leteća besposadna platforma je definisana kao dron, pa će se u tom kontekstu u daljem radu upravo tako i označavati.

Mohammed Mokhtari – Vojna Akademija, Univerzitet odbrane u Beogradu, Veljka Lukića Kurjaka 33, 11042 Beograd, Srbija (e-mail: mokhtari.med91@gmail.com).

Jovan Bajčetić – Vojna Akademija, Univerzitet odbrane u Beogradu, Veljka Lukića Kurjaka 33, 11042 Beograd, Srbija (e-mail: bajce05@gmail.com).

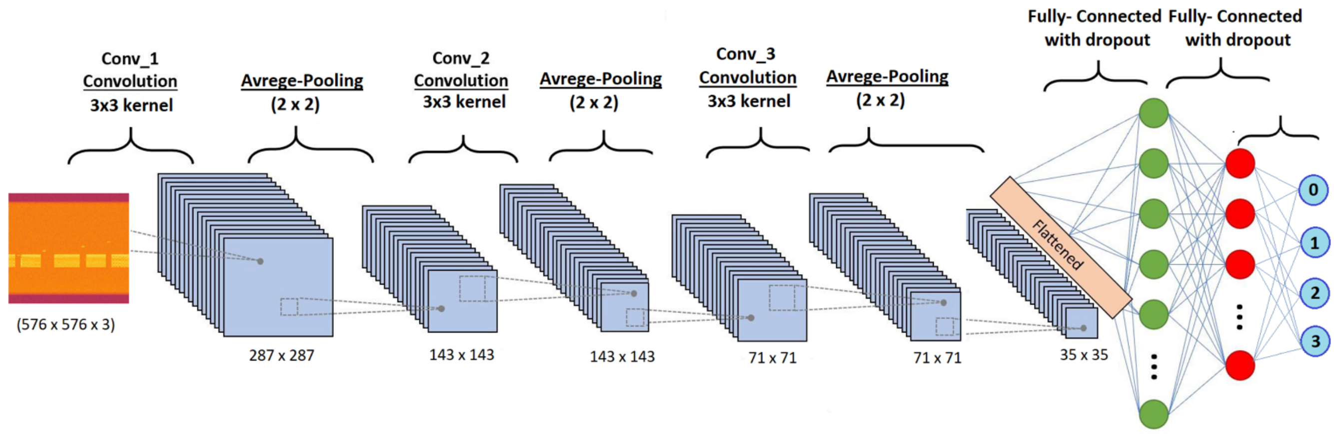
Boban Sazdić-Jotić – Vojna Akademija, Univerzitet odbrane u Beogradu, Veljka Lukića Kurjaka 33, 11042 Beograd, Srbija (e-mail: boban.sazdic.jotic@vs.rs).

U uslovima kada dronovi dobijaju sve više uloga – od snimanja terena, preko prenošenja male količine tereta (dostave), pa sve do mogućnosti prenosa eksplozivnih naprava, nameće se imperativ bolje kontrole ovih letelica koja se može sprovesti kroz kontrolu dela elektromagnetskog spektra u kojem dronovi vrše prenos informacije komandnim i kanalom za prenos slike. Do sada se išlo u više pravaca detekcije i klasifikacije dronova na bazi karakteristika radio emisija koje bi ih izdvojile iz velikog broja ostalih uređaja koji koriste isti frekvencijski opseg:

- Detekcija na bazi MAC (Media Access Control) adrese uređaja [3];
- Detekcija na bazi korišćenih komunikacionih protokola [4];
- Detekcija na bazi „otiska prsta“ radio emisije [5].

Svaka od navedenih metoda primenjena u anti-dron sistemu (ADRO) usled svojih ograničenja i zahteva ima prednosti i nedostatke, pa bi mogućnost implementacije više prikazanih različitih pristupa omogućilo da se dronovi u realnom vremenu detektuju, klasifikuju, lokalizuju i eventualno sprovedu protivmere ukoliko za time postoji potreba. Ovakav pristup zahteva kompromis između pravovremene detekcije, veličine prostornog okvira detekcije, verovatnoće detekcije, kao i kvaliteta i preciznosti upotrebljenih senzora. Do sada se vrlo daleko otišlo u pravcu detekcije i klasifikacije dronova na osnovu karakteristika radio signala u komunikaciji između drona i upravljačke konzole korišćenjem specifičnosti obeležja kompletnog radio saobraćaja, kao jedne u pogledu optimizacije resursa najisplativije metode. U tom pogledu postoji mnogo različitih rešenja korišćenjem metoda dubokog učenja (Deep Learning – DL) koja upotrebljavaju raznovrsne algoritme u zavisnosti od toga za koji problem detekcije i klasifikacije se traži optimalno rešenje – klasifikacija slika [6], 3D objekata [7,8], radio emisija [9, 10, 11], radarskih ciljeva [12], govora i rečenica [13, 14], gasova [15], predviđanje vrednosti akcija [16], itd.

S obzirom na činjenicu da je obučena neuronska mreža utoliko efikasnija u klasifikaciji više različitih radio emisija ukoliko je veća baza podataka pomoću koje se obučava [17], metoda na bazi dubokog učenja oslonjena na snimanjima emisija koji su praktično sprovedeni će biti predstavljena u ovom radu. Da bi se realizovalo istraživanje i razvoj sistema na bazi detekcije i klasifikacije, upotrebljavajući metode dubokog učenja, bilo je potrebno imati na raspolaganju bazu snimaka radio emisija upotrebom kojih bi se izvršio kvalitetan trening i provera (validacija) razvijenog modela koji bi kasnije bio sposoban da sa vrlo visokom verovatnoćom detektuje i klasifikuje dronove na bazi njihove radio emisije.



Sl. 1. Neuronska mreža upotrebljena za detekciju i klasifikaciju dronova (Prva konfiguracija)

Baza snimaka koja je korišćena u ovu svrhu [18] javno je dostupna i sadrži snimke radio emisija dronova u različitim operativnim stanjima, frekvencijskim opsezima (2,4 i 5,8 GHz), kao i snimke radio emisija kada dva i više dronova zajednički deluju na istom prostoru i u istom frekvencijskom opsegu što je čini za sada jedinom te vrste javno dostupnom.

Prikaz istraživanja u ovom radu će se sastojati iz opisa razvijenog modela na bazi dubokog učenja (poglavlje II), prikaza mogućnosti razvijenog modela (poglavlje III), opisa informacionog sistema (poglavlje IV), a na kraju, u okviru zaključka (poglavlje V) će biti rezimirane najbitnije tačke istraživanja i date smernice za buduće istraživanje.

II. Opis modela za detekciju i klasifikaciju

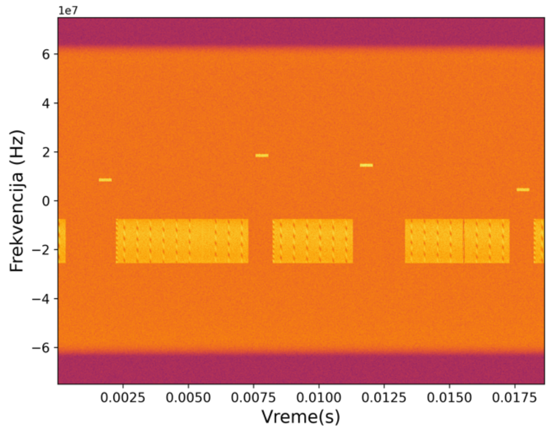
Model za detekciju i klasifikaciju koji je razvijen u okviru istraživanja za početni cilj je imao da obezbedi mogućnost jednostavne detekcije i klasifikacije u okviru četiri klase događaja – nema drona, dron 1, dron 2 i dron 3. S obzirom na uslov nadograđivanja sistema za mogućnost složenije klasifikacije, polazna neuronska mreža je projektovana da bude složenija od toga da zadovolji kriterijum prvog definisanog cilja. Iz tog razloga, razvijeni model sadrži neuronsku mrežu od pet slojeva (Sl. 1.) koja je bila testirana po kriterijumu tačnosti treninga (training accuracy), gubitaka u toku treninga (training loss), provere tačnosti (validation accuracy) i gubitaka u toku provere tačnosti (validation loss) za različite slučajeve kompleksnosti ulaznih podataka. Neuronska mreža korišćena u istraživanju pripada tipu konvolucione neuronske mreže (Convolutional Neural Network – CNN), kao ulazne podatke koristi 2D slike spektrograma i predstavlja adaptaciju funkcionalne mreže sa strukturom za klasifikaciju audio signala [19]. Optimizacija mreže za slučaj klasifikacije koji je naveden svodio se na pronalaženje optimalnog odnosa između kompleksnosti i veličine ulaznih podataka (slika), verovatnoće detekcije i klasifikacije i vremena potrebnog za treniranje modela neuronske mreže. Shodno tome, vršena je optimizacija na osnovu broja odbiraka sirovog snimka i na taj način vršen uticaj na rezoluciju spektograma i dimenzije njegove slike. Uticaj rezolucije slike na performanse CNN je proučavan u literaturi [20], kao i arhitektura jednodimenzionalne konvolucione neuronske mreže (1DCNN) za detekciju i klasifikaciju UAV [21].

Formirana mreža je prilagođavana formatu ulaznih podataka (adaptirani su slojevi posledično veličini podataka koji se koriste) čije su vrednosti dobijane segmentacijom sirovih podataka iz navedene baze snimljenih radio emisija sa dronova, a vrednosti dobijene nakon segmentacije su prikazane u Tabeli I. Za svaku od četiri moguće konfiguracije je biran adekvatan broj odbiraka po segmentu čime je definisano trajanje svakog segmenta, kao i ukupan broj segmenata proistekao iz sirovog snimka. Raspodela ukupnog broja segmenata po klasama (Nema drona, Dron 1, Dron 2 i Dron 3) je predstavljena u donjem delu Tabele I. Eksperiment se svodio na četiri različita slučaja (konfiguracije), a dronovi 1, 2 i 3 su bili, respektabilno DJI Phantom IV, DJI Mavic 2 Zoom i DJI Mavic 2 Enterprise.

TABELA I
Ulazni podaci – segmentacija

Konfiguracija	Broj odbiraka po segmentu	Trajanje segmenta (ms)	Ukupan broj segmenata
Prva	2796128	18.64	360
Druga	1398064	9.32	720
Treća	699032	4.66	1440
Četvrta	200319	1.34	5025
Nema drona	Dron 1	Dron 2	Dron 3
72	96	96	96
144	192	192	192
288	384	384	384
1005	1340	1340	1340

Od dobijenih segmentiranih podataka su nakon izračunavanja Furijeove transformacije (Hanning tip prozora, 2048 tačaka za FFT, 128 tačaka preklapanja, amplituda predstavljena u dB) kreirane slike spektrograma u .png formatu (Sl. 2.) prilagođene za ulaz u svaku od konfiguracija neuronske mreže (Tabela II).



Sl. 2. Primer izgleda spektrograma

TABELA II
KARAKTERISTIKE SLIKA SPEKTROGRAMA

Konfiguracija	Rezultujuća matrica spektrograma	Rezultujuća slika (pixels)
Prva	2048 x 1456	576 x 576
Druga	2048 x 727	406 x 406
Treća	2048 x 364	288 x 288
Četvrta	2048 x 104	154 x 154

Od ukupnog broja segmentiranih podataka, 70 % je iskorišćeno za treniranje modela, a 30 % za proveru. Raspodela broja slika po skupovima, za svaku konfiguraciju je predstavljena u Tabelama III i IV. Tabela III predstavlja raspodelu slika po četiri definisane klase (Nema drona, Dron 1, Dron 2, Dron 3) za sve četiri trenirane konfiguracije. Tabela IV prikazuje raspodelu broja slika za iste četiri klase, ali za proveru tačnosti istreniranih modela za sve četiri konfiguracije.

TABELA III
PARAMETRI ZA TRENIRANJE MODELA

Konfiguracija	Ukupno slika	Nema drona	Dron 1	Dron 2	Dron 3
Prva	251	50	67	67	67
Druga	502	100	134	134	134
Treća	1005	201	268	268	268
Četvrta	3514	703	937	937	937

TABELA IV
PARAMETRI ZA PROVERU MODELA

Konfiguracija	Ukupno slika	Nema drona	Dron 1	Dron 2	Dron 3
Prva	109	22	29	29	29
Druga	218	44	58	58	58
Treća	435	87	116	116	116
Četvrta	1511	302	403	403	403

III. MOGUĆNOSTI MODELA ZA DETEKCIJU I KLASIFIKACIJU

Treniranje modela dubokog učenja izvršeno je po već navedenim podacima u ukupno 30 epoha za svaku od konfiguracija. Rezultati neimenovane jedinice tačnosti treninga (training accuracy) čija vrednost može biti od 0 do 1, a koji predstavlja odnos između broja tačnih predikcija i ukupnog broja predikcija u funkciji od broja epoha, za sve četiri konfiguracije dati su na Sl. 3. Neimenovana jedinica gubitaka u toku treninga (training loss) u funkciji od broja epoha čija vrednost može pripadati skupu $[0,\infty)$, a čija manja vrednost implicira bolju predikciju predstavljeni su na Sl. 4.

Krajnja tačnost treninga, kao i gubitaka u toku treninga za svaku od konfiguracija su predstavljeni u Tabeli V.

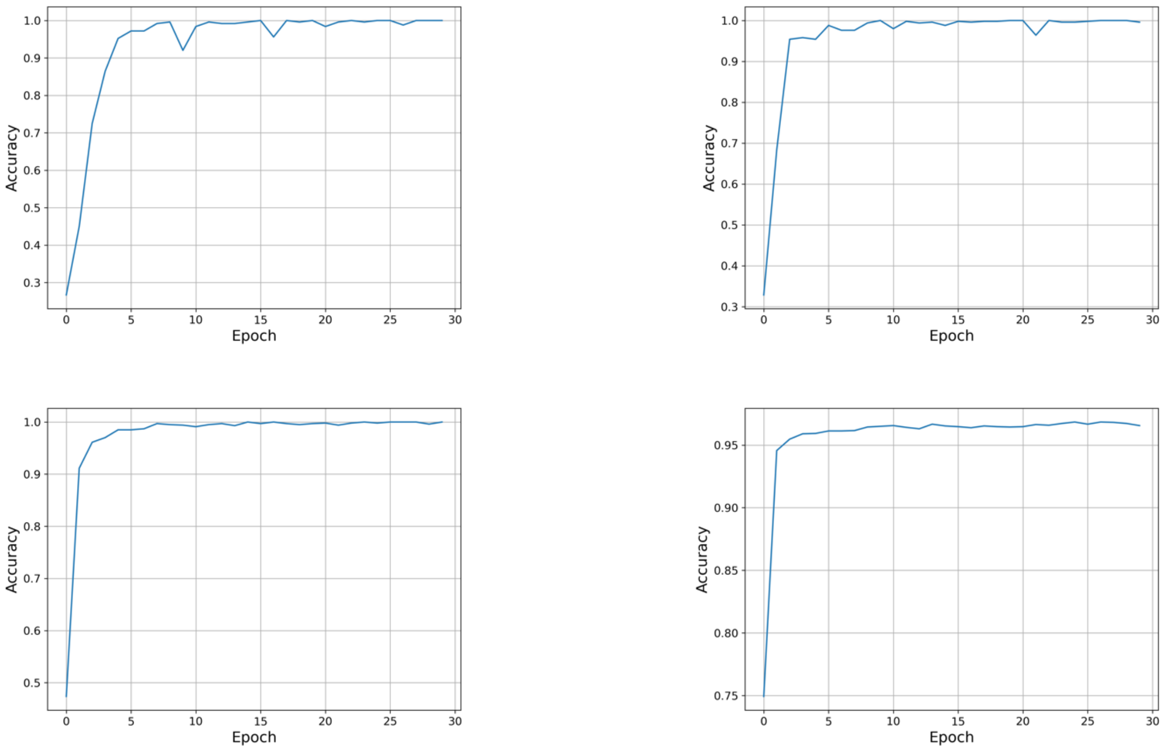
TABELA V
VREDNOSTI TRENIRANJA ZA SVAKU KONFIGURACIJU

Konf.	Broj epoha	Tačnost treninga	Gubici u toku treninga	Provera tačnosti	Gubici u toku provere tačnosti
Prva	30	1.00	$2.64\cdot10^{-6}$	1.00	$1.01\cdot10^{-5}$
Druga	30	1.00	$6.23\cdot10^{-4}$	0.99	0.01
Treća	30	1.00	$2.65\cdot10^{-6}$	1.00	$2.23\cdot10^{-6}$
Četvrta	30	0.97	0.11	0.97	0.12

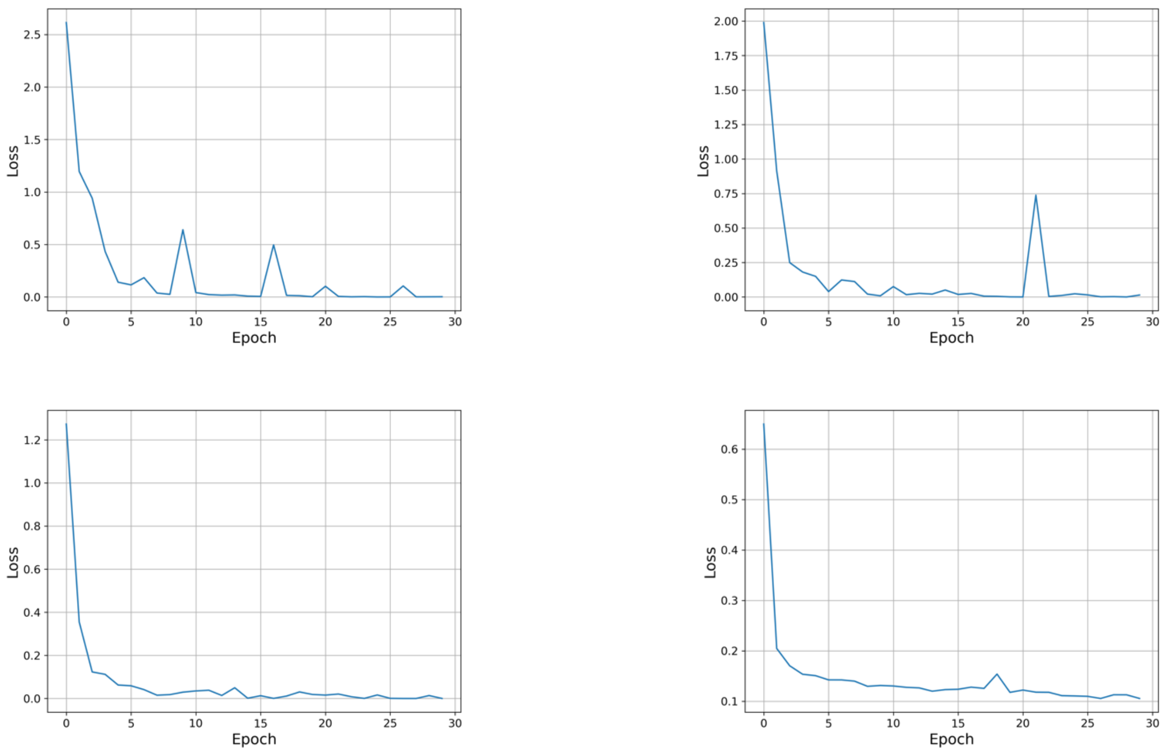
Na osnovu prikazanog se može zaključiti:

1. Uzimajući u obzir kriterijum tačnosti treninga, prve tri konfiguracije omogućavaju klasifikaciju sa najvišom mogućom verovatnoćom, kao i sa vrlo malim gubicima u toku treninga (reda veličine 10^{-5}) što se može videti u Tabeli V;
2. U pogledu provere tačnosti, kao i u pogledu gubitaka u toku provere tačnosti, prva i treća konfiguracija daju dobre i približno slične rezultate.
3. Četvrta konfiguracija je po pitanju sva četiri kriterijuma lošija od prve tri, sa krajnjom tačnošću treninga i provere tačnosti od 0,97.
4. Model (konfiguracija) koji bi predstavljao kompromis između brzine treniranja, tačnosti i kompleksnosti (veličine) ulaznih podataka i na taj način bio optimalan za korišćenje u informacionom sistemu koji bi funkcionisao u realnom vremenu bi bio Treći model.

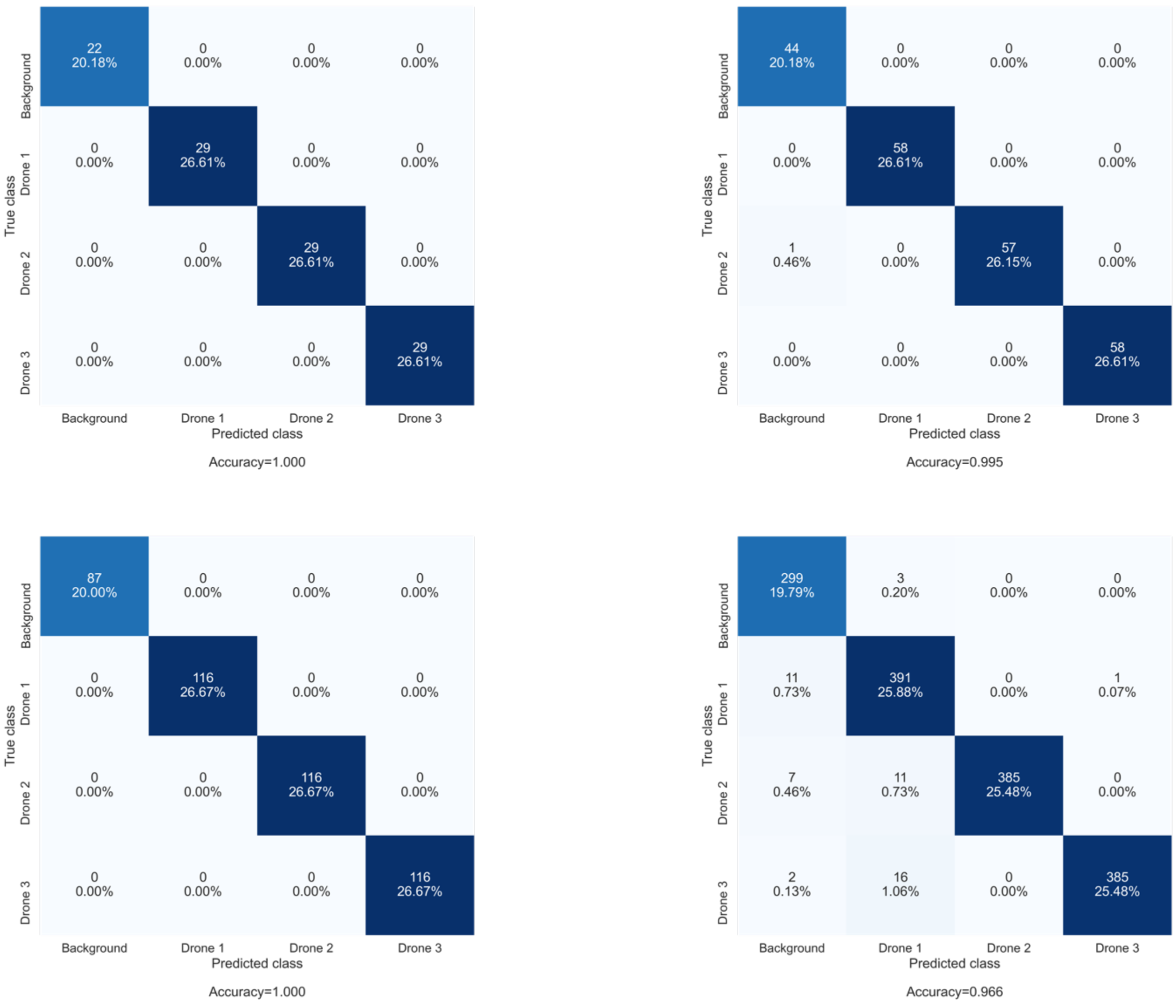
Matrice konfuzije na osnovu kojih se mogu ustanoviti kompletne performanse sve četiri konfiguracije, predstavljene su na Sl. 5. Ove matrice predstavljaju grafički prikaz klasifikacije po broju slika i njihovom procentualnom udelu u ukupnom broju slika koje su svrstane u određenu kategoriju (klasu). Ova klasifikacija može biti pozitivna (tačna) ili negativna (netačna), pa se na osnovu toga može zaključiti da je konfiguracija 4 najviše grešaka napravila prilikom klasifikacije emisije Dron 1 koju je svrstala u kategoriju Dron 3.



Sl. 3. Tačnost treninga (training accurracy) za konfiguracije: 1 (gore levo), 2 (gore desno), 3 (dole levo) i 4 (dole desno)



Sl. 4. Gubici u toku treninga (training loss) za konfiguracije: 1 (gore levo), 2 (gore desno), 3 (dole levo) i 4 (dole desno)



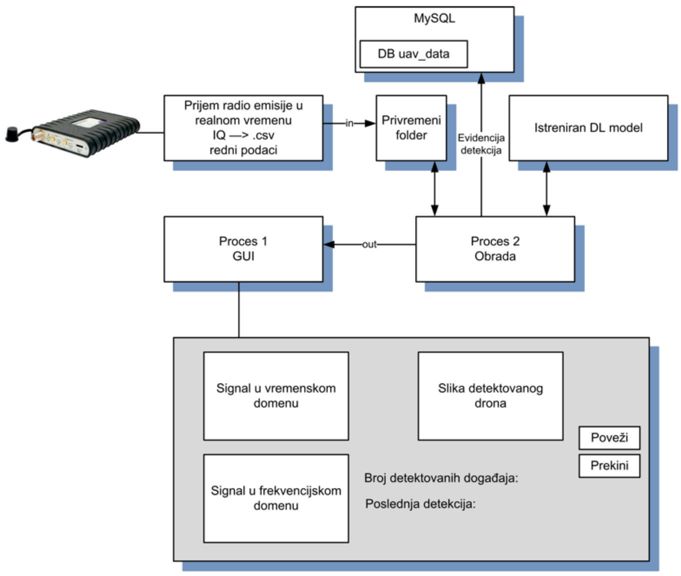
Sl. 5. Matrice konfuzije (confusion matrix) za konfiguracije: 1 (gore levo), 2 (gore desno), 3 (dole levo) i 4 (dole desno)

IV. INFORMACIONI SISTEM ZA DETEKCIJU I KLASIFIKACIJU

Produkt predstavljenog istraživanja je operativni informacioni sistem koji omogućava detekciju i klasifikaciju dronova u realnom vremenu. Ovaj jednostavni sistem je razvijen u Python programskom okruženju, proširiv je i nadogradiv za buduću upotrebu i uvezivanje sa ostalim komponentama, u skladu sa potrebama. Blok šema informacionog sistema predstavljena je na Sl. 6.

Akvizicija podataka u realnom vremenu realizuje se analizatorom spektra Tektronix RSA306B USB, a pomoću programa koji priprema i segmentuje podatke za upotrebu u Procesu 2 aplikacije – Obrada. Proces obrade koristi istreniran DL model konfiguracije 3 i u toku rada, paralelno sa prikazom detekcije u GUI (Graphic User Interface) procesu, vrši upis detekcija u MySQL bazu radi evidencije i kasnije analize.

Ovaj sistem predstavlja osnovu za razvoj kompleksnijeg informacionog sistema koji će u skladu sa klasom detekcije, vršiti određene protivmere radi zaštite ili efikasnije upotrebe dostupnog elektromagnetskog spektra u slučaju koegzistencije više besposadnih niskoletećih platformi na istom području.



Sl. 6. Blok šema informacionog sistema za detekciju i klasifikaciju niskoletećih besposadnih platformi (SDKNBP)

V. Zaključak

Istraživanjem predstavljenom u ovom radu došlo se do formiranja optimalnog modela dubokog učenja sposobnog da sa maksimalnom mogućom verovatnoćom klasifikuje detekcije emisija sa dronova u četiri moguće klase. Optimizacijom ulaznih parametara po strukturi i veličini, omogućena je detekcija i klasifikacija u realnom vremenu upotrebom „Real-Time" analizatora spektra. Informacioni sistem koji je razvijen na bazi optimalnog DL modela ima mogućnost prikaza i zapisa detektovanih događaja, kao i mogućnost nadgradnje radi daljeg razvoja u željenom smeru.

Predstavljene konfiguracije predstavljaju adaptacije neuronske mreže od pet slojeva koja će u budućem istraživanju biti proširena i primenjena u razvijenom informacionom sistemu za mogućnost klasifikacije tri vrste drona u pet različitih scenarija upotrebe – nema drona (pozadinski šum), dron je povezan sa konzolom za upravljanje, dron lebdi, dron leti, dron leti i prenosi sliku.

Zahvalnica

Ovo istraživanje je izvedeno u okviru projekta VA-TT/3/20-22 koje finansira Univerzitet odbrane u Beogradu.

Literatura

[1] E. Mitka and S. G. Mouroutsos, "Classification of Drones," *Am. J. Enginnering Res.*, vol. 6, no. 7, pp. 36–41, 2017.

[2] Altigator, "Drone, UAV, UAS, RPA or RPAS." [Online]. Available: https://altigator.com/drone-uav-uas-rpa-or-rpas/. [Accessed: 25-Jun-2021].

[3] Peacock, M., & Johnstone, M. N. (2013). Towards detection and control of civilian Unmanned Aerial Vehicles. *Australian Information Warfare and Security Conference*, *June*, 1–8. https://doi.org/10.4225/75/57a847dfbefb5

[4] D-Fend. (2021). *D-Fend Solutions A.D. Ltd.* https://www.d-fendsolutions.com/

[5] Nguyen, P., Ravindranatha, M., Nguyen, A., Han, R., & Vu, T. (2016). Investigating cost-effective RF-based detection of drones. *Proceedings of the 2nd Workshop on Micro Aerial Vehicle Networks, Systems, and Applications for Civilian Use*, *October 2018*, 17–22. https://doi.org/10.1145/2935620.2935632

[6] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1145/3065386

[7] Nair, V., & Hinton, G. E. (2009). 3D Object Recognition with Deep Belief Nets. *Advances in Neural Information Processing Systems 22 - Proceedings of the 2009 Conference*, *22*, 1339–1347. https://proceedings.neurips.cc/paper/2009/file/6e7b33fdea3adc80ebd648fffb665bb8-Paper.pdf

[8] Pathak, A. R., Pandey, M., & Rautaray, S. (2018). Application of Deep Learning for object detection. *Procedia Computer Science*, *132*, 1706–1717. https://doi.org/10.1016/j.procs.2018.05.144

[9] Peng, S., Jiang, H., Wang, H., Alwageed, H., Zhou, Y., Sebdani, M. M., & Yao, Y.-D. (2019). Modulation classification based on signal constellation diagrams and Deep Learning. *IEEE Transactions on Neural Networks and Learning Systems*, *30*(3), 718–727. https://doi.org/10.1109/TNNLS.2018.2850703

[10] Zha, X., Peng, H., Qin, X., Li, G., & Yang, S. (2019). A Deep Learning framework for signal detection and modulation classification. *Sensors*, *19*(18), 4042. https://doi.org/10.3390/s19184042

[11] Zhou, S., Yin, Z., Wu, Z., Chen, Y., Zhao, N., & Yang, Z. (2019). A robust modulation classification method using convolutional neural networks. *EURASIP Journal on Advances in Signal Processing*, *2019*(1), 21. https://doi.org/10.1186/s13634-019-0616-6

[12] Alhadhrami, E., Al-Mufti, M., Taha, B., & Werghi, N. (2019). Learned micro-Doppler representations for targets classification based on spectrogram images. *IEEE Access*, *7*, 139377–139387. https://doi.org/10.1109/ACCESS.2019.2943567

[13] Karita, S., Watanabe, S., Iwata, T., Ogawa, A., & Delcroix, M. (2018). Semi-Supervised End-to-End Speech Recognition. *Interspeech 2018*, 2–6. https://doi.org/10.21437/Interspeech.2018-1746

[14] Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1746–1751. https://doi.org/10.3115/v1/D14-1181

[15] Narkhede, P., Walambe, R., Mandaokar, S., Chandel, P., Kotecha, K., & Ghinea, G. (2021). Gas detection and identification using multimodal artificial intelligence based sensor fusion. *Applied System Innovation*, *4*(1), 1–14. https://doi.org/10.3390/asi4010003

[16] Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques. *Expert Systems with Applications*, *42*(1), 259–268. https://doi.org/10.1016/j.eswa.2014.07.040

[17] Mohammad F. Al-Sa'd, Abdulla Al-Ali, Amr Mohamed, Tamer Khattab, Aiman Erbad, (2019). RF-based drone detection and identification using deep learning approaches: An initiative towards a large open source drone database, *Future Generation Computer Systems*, 100, 86-97, ISSN 0167-739X, https://doi.org/10.1016/j.future.2019.05.007

[18] Sazdić-Jotić, B. M., Pokrajac, I., Bajčetić, J., Bondžulić, B. P., Joksimović, V., Šević, T., & Obradović, D. (2020). *VTI_DroneSET_FFT*. Mendeley Data. https://doi.org/10.17632/s6tgnnp5n2.1

[19] https://github.com/nageshsinghc4/Audio-Data-Analysis-Using-Deep-Learning

[20] Kannojia, S.P., & Jaiswal, G. (2018). Effects of Varying Resolution on Performance of CNN based Image Classification An Experimental Study, *International Journal of Computer Sciences and Engineering,* 6 (9), 451-456, E-ISSN 2347-2693, https://doi.org/10.26438/ijcse/v6i9.451456

[21] Al-Emadi, S., Al-Senaid, F. (2020). Drone Detection Approach Based on Radio-Frequency Using Convolutional Neural Network, *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, 29-34, https://doi.org/10.1109/ICIoT48696.2020.9089489

Abstract

The research which is described in this paper presents the potentials of the Drone Detection and Classification System (DDCS) proposed concept which is based on Deep Learning (DL) methods. The aim of the project is the development of the operational system which would be capable of real time drone detection and classification based on their radio emission characteristics. Detection and classification methods used in this research are built with developed convolutional neural network which is trained using the formed radio emission database consisted of recordings originated from several commercially available drones. The initial testing of the research results shows 100 % detection and classification probability for four independent classes – no drone, drone 1, drone 2 and drone 3 which is a good basis for a future more detailed classification development.

Drone Detection and Classification System (DDCS)

Mohammed Mokhtari, Jovan Bajcetic and Boban Sazdic-Jotic

Evaluacija dometa LoRa IoT primopredajnika u urbanom i ruralnom okruženju

Dejan N. Milić, *Member, IEEE*, Slavimir Stošović, Dejan Stevanović i Jelena Anastasov, *Student Member, IEEE*

***Apstrakt*—U ovom radu su prikazani praktični rezultati merenja nivoa signala prilikom prenosa podataka korišćenjem platforme tipa LoRa. Iako je ciljna primena ovakvih platformi na prvom mestu internet inteligentnih objekata u urbanim okruženjima, u radu je razmotrena i mogućnost njihove primene u ruralnim okruženjima. Jedna od ideja za primenu tehnologije jeste praćenje živog sveta u okruženjima koja se nalaze van naseljenih oblasti. Domet radio veze u ovakvom okruženju može biti znatno veći nego u urbanoj sredini, zavisno od uslova optičke vidljivosti i prepreka koje se nalaze na liniji između predajnika i prijemnika. Dobijeni rezultati pokazuju da se pri povoljnim uslovima zadovoljavajući rezultati prenosa mogu dobiti i na udaljenostima većim od 10 km.**

***Ključne reči*—prostiranje RF signala; LoRa; IoT, feding; efekat senke.**

I. UVOD

INTERNET inteligentnih objekata (*Internet of Things* - IoT) predstavlja koncept mreže za prenos podataka unutar koje su objekti opremljeni senzorima, softverom i uopšte – neophodnom tehnologijom u cilju razmene podataka između uređaja i sistema, preko interneta [1], [2]. Ovakav koncept je ponikao iz konvergencije više različitih tehnologija, pri čemu je ključna sposobnost komunikacija između objekata unutar mreže, kao i objekata i infrastrukture. Ove sposobnosti mogu uspešno omogućiti različite tehnologije bežičnih komunikacija [3], od kojih najvažnije spadaju u tri kategorije prema svom dometu: tehnologije kratkog, srednjeg i većeg dometa. U tehnologije kratkog dometa se ubrajaju Bluetooth, NFC, RFID, Wi-Fi, ZigBee i slične tehnologije koje imaju domet prenosa do desetak metara udaljenosti. U kategoriju tehnologija srednjeg dometa spadaju npr. LTE ili 5G, obzirom da omogućavaju pristup u okviru nekoliko stotina metara, odnosno do najbliže bazne stanice. Sa druge strane, tehnologije većeg dometa mogu biti npr. LoRa, [4], NB-IoT i slične, koje mogu omogućiti domet u krugu od nekoliko kilometara.

U ovom radu smo eksperimentalno testirali prenos LoRa modula namenjenih za IoT sisteme većih dometa. Cilj testiranja jeste okvirna procena dometa ovakvih uređaja, uzimajući u obzir karakteristike prenosnog kanala [5]-[7], u dva scenarija korišćenja:

- u urbanim uslovima, odnosno uslovima koji podrazumevaju prostiranje u urbanim sredinama i zatvorenom prostoru, i
- u ruralnim uslovima, odnosno uslovima koji se odnose na otvoreni prostor i optičku vidljivost u širokom opsegu pokrivanja.

Ove informacije su neophodne kako bi se za odgovarajuću konkretnu primenu odabrala neophodna tehnologija bežičnog pristupa i sagledala izvodljivost zacrtanih ciljeva.

Struktura ovog rada je seledeća: nakon uvoda, sledi odeljak II u kome su opisani relevantni delovi korišćenog hardvera. U odeljku III predstavljena su merenja koja se odnose na evaluaciju antene, dok su odeljci IV i V posvećeni rezultatima dobijenim u urbanoj i ruralnoj sredini, respektivno. Zaključak rada sažeto je dat u poslednjem odeljku VI, nakon koga sledi zahvalnica i spisak korišćene literature.

II. HARDVER

Hardver koji je korišćen kao test platforma je osnova projekta CANANDI, čiji je cilj realizacija komercijalnog sistema namenjenog praćenju lovačkih pasa u ruralnoj sredini. Canandi sistem se sastoji od ogrlice koju nosi pas, prijemnika koji se nalazi kod vlasnika psa i mobilne aplikacije. Komunikacija između ogrlice psa i prijemnika koji se nalazi kod vlasnika zasnovana je na tehnologiji koja podržava LoRa standard. Canandi pomaže vlasnicima da u svakom trenutku znaju lokaciju njihovih pasa, pri čemu predstavlja inovativno rešenje na tržištu jer sadrži nekoliko jedinstvenih odlika:

- Dvostruki način prenosa signala (LoRa, GSM) za povećanje pouzdanosti sistema
- IoT komunikaciona mreža između lovaca na terenu koja omogućava deljenje i korišćenje podataka za zajedničko praćenje izgubljenog psa
- Mašinsko učenje u analizi prikupljenih podataka i davanje korisnih saveta lovcima za unapređenje obuke pasa.

Primopredajnik je u osnovi komercijalno dostupan pod oznakom RFM95W, a prototip koji je korišćen u testu je realizovan sa pomenutim modulom primopredajnika. Predajni modul sadrži izlazni pojačavač snage koji je projektovan po kriterijumu optimalne efikasnosti, pri čemu tipično može da obezbedi snagu od 25 mW, odnosno +14 dBm na izlazu, kada je izlazno opterećenje realna otpornost od 50 Ω. Pored

Dejan N. Milić – Elektronski fakultet, Univerzitet u Nišu, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: dejan.milic@ elfak.ni.ac.rs).
Slavimir Stošović – Akademija tehničko-vaspitačkih strukovnih studija Niš, 1800 Niš, Srbija (e-mail: slavimir.stosovic@akademijanis.edu.rs).
Dejan Stevanović, Elektronski fakultet, Univerzitet u Nišu, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: stevanovic_dejan@yahoo.com).
Jelena Anastasov – Elektronski fakultet, Univerzitet u Nišu, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: jelena.anastasov@ elfak.ni.ac.rs).

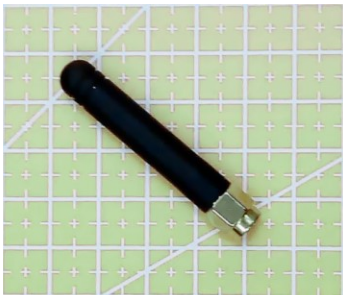
standardnih formata modulacije zasnovanih na digitalnoj amplitudskoj i frekvencijskoj modulaciji - OOK, FSK, GFSK i GMSK, koji obezbeđuju kompatibilnost sa postojećim standardima - IEEE802.15.4g i bežičnim MBUS standardom, modul ima i podršku za LoRa modulaciju/demodulaciju. Modulacija koju koristi LoRa se zasniva na prenosu signala sa proširenim spektrom, pri čemu LoRa predstavlja derivativ tehnike proširenog spektra sa frekvencijskim čirpom (CSS). Prednosti koju ovakva vrsta modulacija ima nad standardnim digitalnim modulacijama se ogledaju u boljoj osetljivosti prijemnika, koja može biti i ispod nivoa šuma ulaznog stepena, zatim u boljoj selektivnosti i smanjenju kanalne interferencije, kao i u boljim karakteristikama u pogledu efekta fedinga.

III. ANTENA

Antena predstavlja osnovu gotovo svakog IoT projekta koji koristi bežične komunikacije. Sistem bežičnog radio prenosa suštinski zavisi od antene kao interfejsa koji je nepohodan kako bi se radio talasi efikasno emitovali u prostor izvan predajnika. Uobičajeno je da se sistem prenosa, uključujući antenu, projektuje na početku, a da se nakon toga dodaju ostali elementi sistema.

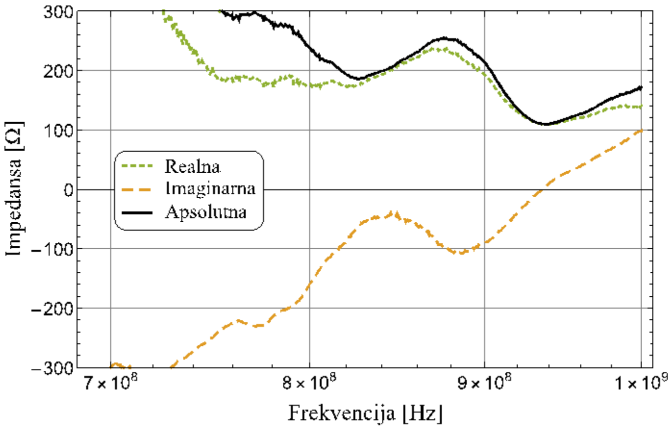
U konkretnom slučaju realizovanog prototipa antena je zamišljena kao eksterni element, a štampana ploča je realizovana tako da linija koja vodi od izlaznog stepena do konektora antene bude veoma kratka. S obzirom da u samom početku nije obavljen izbor konkretne antene, već je ostavljena sposobnost testiranja različitih mogućnosti, izlazni port antene je zamišljen kao U/FL konektor na koji je moguće priključiti različite antene radi daljeg testiranja. Sa druge strane, konkretna namena uređaja ima određena ograničenja u pogledu dimenzija, tako da je prototip korišćen prilikom testiranja umesto direktnog montiranja antene imao opciju napojnog voda koji sa jedne strane ima U/FL konektor, a sa druge strane SMA konektor koji se montira na plastično kućište uređaja.

Antena koja je korišćena u ovom testu je monopol antena niske nabavne cene, sa ravnim SMA konektorom, prikazana na slici 1.



Sl. 1. Eksterna antena uređaja.

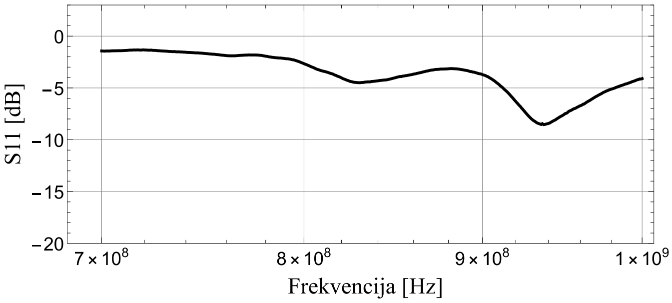
Obzirom da monopol antena zahteva ravan uzemljenja koja u ovom prototipu nije predviđena, intuitivno se može pretpostaviti da njene karakteristike nisu optimalne. Zbog toga



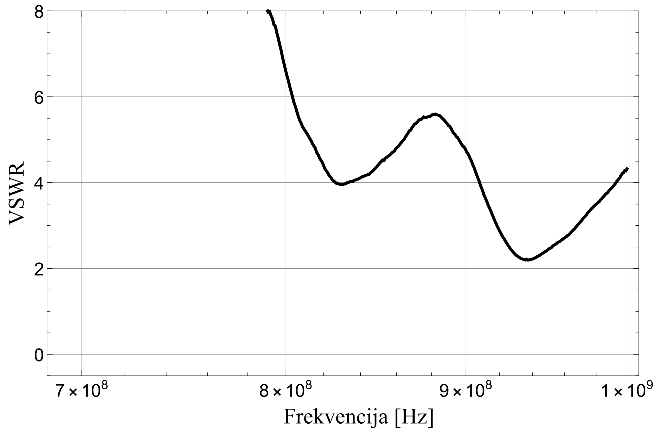
Sl. 2. Realni i imaginarni deo imedanse antene, izmereni u odsustvu ravni uzemljenja.

smo najpre obavili merenje impedanse ove antene u konfiguraciji sa koaksijalnim napajanjem, bez postojanja ravni uzemljenja. Dobijeni rezultati su prikazani na slici 2. Sa slike se vidi da na frekvenciji 868 MHz impedansa antene iznosi oko 250 Ω. Impedansa je primarno realna, dok imaginarni deo impedanse iznosi oko -100 Ωi.

Kako u realizaciji projekta još uvek nije donesena odluka o korišćenju konkretne antene, dalji koraci ka prilagođenju antene nisu preduzimani i testiranje prototipa je nastavljeno bez implementacije kola za prilagođenje. Jedan od faktora koji doprinosi ovakvoj odluci jeste i koaksijano napajanje antene



a.)



b.)

Sl. 3. a) Koeficijent refleksije i b) naponski koeficijent stojećih talasa.

putem RG174 kabla čija dužina nije precizno kontrolisana, tako da bi prilagođenje moralo da se posebno projektuje za svaki od uređaja na osnovu izmerenih vrednosti impedanse. Imajući ove činjenice u vidu, izmerene su karakteristike refleksije za konkretni slučaj i rezultati su prikazani na slici 3.

Slika 3. pokazuje da koeficijent refleksije, S_{11}, gledano od strane SMA konektora iznosi oko -3 dB, što odgovara naponskom koeficijentu stojećih talasa od 5.5. Iako ove vrednosti nisu blizu optimalnih, iz pomenutih razloga kolo za prilagođenje nije dalje razmatrano, već su performanse testirane sa datim prototipom uz komentar da će njihovo poboljšanje definitivno biti moguće u daljoj realizacjii projekta.

IV. Test u urbanoj sredini

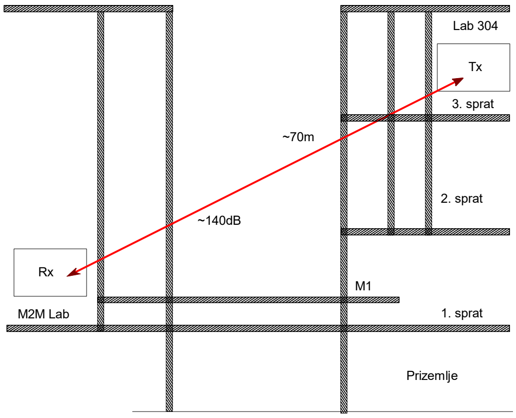
Iako su IoT uređaji koji sadrže LoRa module obično deklarisani kao uređaji dugog dometa, u urbanim sredinama ovakvu tvrdnju treba prihvatiti sa izvesnom rezervom. Pri prostiranju u urbanoj sredini, radio signal podleže različitim propagacionim efektima od kojih su dominantni [8], [9]:

- efekat fedinga, i
- efekat senke

Feding nastaje kao posledica interfenecije više talasa koji se prostiru različitim mogućim putanjama između predajnika i prijemnika. Obzirom da u urbanim sredinama postoji veći broj ravnih/neravnih površina koje mogu biti reflektivne ili predstavljati centre rasejanja radio talasa, na mestu prijema će doći do suštinskog kombinovanja velikog broja komponenti signala, koje imaju različite amplitude i faze. Kombinovanje može biti konstuktivno ili destruktivno, zavisno od konkretne raspodele faza i kašnjenja signala [8]. U vezi sa različitim raspodelama komponenti signala, mogu se formulisati različiti modeli fedinga koji na statistički način tretiraju ove propagacione efekte [10].

U testu koji se odnosi na urbanu sredinu, odlučili smo da predajnik i prijemnik udaljimo do te mere da signali na prijemu budu iznad nivoa šuma, kako bismo mogli da ih detektujemo sa relativno velikom verovatnoćom, odnosno tako da verovatnoća otkaza prenosa bude niska. Ovakvu situaciju smo postigli postavljanjem predajnika u laboratoriju 304 na Elektronskom fakultetu, dok je prijemnik bio smešten u suprotnom delu zgrade, u prostorije M2M laboratorije koja je trenutno u procesu osnivanja. Šematski prikaz pozicija predajnika i prijemnika, kao i fotografija iz vazduha na kojoj su označene ove pozicije, predstavljene su na slici 4.

Udaljenost predajnika i prijemnika je oko 70 m, pri čemu je ukupno slabljenje na ovoj deonici oko 140 dBm. Obzirom da se slabljenje u slobodnom prostoru može proceniti na 68 dB pod pretpostavkom izotropne antene, ostatak slabljenja se mora pripisati slabljenju na zidovima na putu prostiranja, što iznosi oko 72 dB. Sa slike 4.a se vidi da se na putu prostiranja nalazi 5 horizontalnih zidova i jedna podna površina, pa se može proceniti da je slabljenje na individualnom betonskom zidu/podu oko 12 dB.

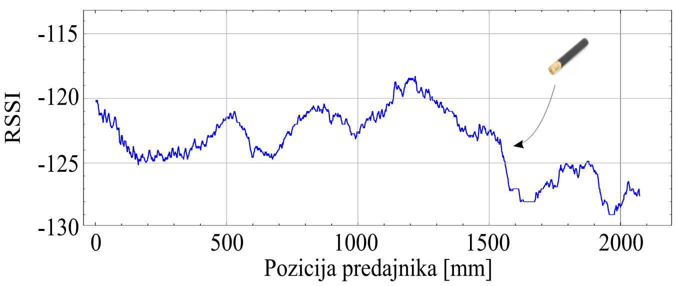


a.)

b.)

Sl. 4. Šematski prikaz lokacija predajnika i prijemnika unutar zgrade Elektronskog fakulteta prilikom procene propagacionih efekata za slučaj urbane sredine.

Konkretna realizacija efekta fedinga je proučena za ovaj slučaj korišćenjem mehanizma za precizno pozicioniranje predajnika u laboratoriji 304. Prilikom pomeranja predajnika duž linijske putanje sa korakom od 1 mm, na prijemu je svaki put zabeležen nivo primljenog signala. Procena primljenog nivoa se oslanja na informaciju dobijenu od samog prijemnika, a ova informacija se beleži u obliku indikatora jačine primljenog signala (RSSI), koji približno odgovara primljenoj snazi u jedinicama dBm. Na slici 5. je prikazan skup prikupljenih podataka dobijenih skeniranem putanje predajnika u dužini od 2.1 m, koji sadrži ukupno 2100 tačaka. Potrebno je pomenuti da ispod indikatora -129 nije primljen ni jedan paket, odnosno svi primljeni paketi su iznad ovog nivoa.



Sl. 5. Indikator jačine signala na prijemu u zavisnosti od precizne pozicije predajnika, za slučaj prostiranja u urbanoj sredini.

Ovde treba navesti da su parametri komunikacije podešeni na sledeći način: propusni opseg – 62,5 kHz, faktor proširenja spektra – 10, kodni faktor – 4/5, što odgovara bitskoj brzini od 488 b/s. Naravno, niska bitska brzina odgovara IoT namenama koje prenose malu količinu podataka, zauzimajući zadati kanal u toku veoma kratkog vremenskog intervala i sa relativno velikom pauzom između sukcesivnih transmisija.

Svakako, treba imati u vidu i prirodu modulacije sa proširenim spektrom, koja obezbeđuje izvesni nivo imunosti na efekat fedinga u odnosu na uskopojasne signale. Obzirom da je ukupna varijacija signala usled efekta fedinga iznosila oko +/-5 dB, očigledno je da u urbanim uslovima, kada ne postoji linija direktne optičke vidljivosti između predajnika i prijemnika, osnovni problem sistema predstavlja efekat senke oličen u slabljenju na fizičkim preprekama.

V. Test na otvorenom prostoru

Kada se razmatraju performanse IoT sistema zasnovanog na LoRa primopredajnicima u ruralnim uslovima, treba uočiti da su tada propagacioni efekti značajno različiti od uslova u urbanim sredinama. Snaga signala na prijemu se usled prostiranja u slobodnom prostoru može proceniti shodno ilustraciji na slici 6. i poznatim parametrima koji se odnose na izotropnu antenu, kao [6]:

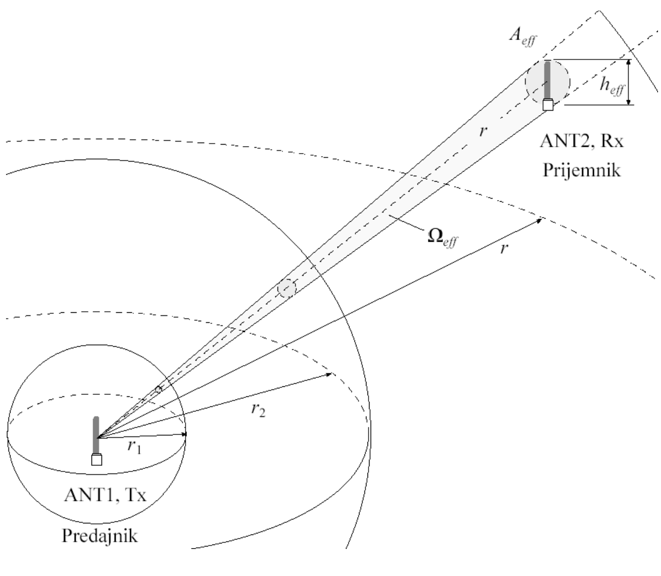
$$\left[P_{Rx}\right]_{dB} = \left[P_{Tx}\right]_{dB} + 2\left(10\log_{10}\frac{c}{4\pi f_0} - 10\log_{10} r\right) \quad (1)$$

Analizom ove jednačine može se utvrditi da slabljenje koje odgovara nivou osetljivosti prijemnika zahteva ekstremno velike udaljenosti. Ova činjenica ima i praktičnu potvrdu u merenjima koja su pokazala da je u povoljnim uslovima pri kojima postoji linija optičke vidljivosti, prijem LoRa signala moguć i na udaljenostima većim od 700 km. U pomenutom slučaju je optička vidljivost obezbeđena time što je predajnik podignut na visinu od 38 km uz pomoć meteorološkog balona.

Jedna od osnovnih pretpostavki je da u ruralnim uslovima može postojati optička vidljivost između predajnika i prijemnika. Ipak, ovu pretpostavku treba uzeti sa rezervom jer je upadljivo zavisna od konfiguracije terena. Čak i kada pretpostavimo da na površini Zemlje ne postoje uzvišenja i udubljenja, odnosno kada Zemljinu površinu posmatramo kao savršenu sferu, jednostavnim geometrijskim razmatranjem možemo doći do formule koja opisuje maksimalnu daljinu optičke vidljivosti u zavisnosti od visine na kojoj je postavljena antena. Ova formula se može aproksimirati izrazom:

$$d_{[km]} \approx 3.57\sqrt{h_{[m]}}\,, \quad (2)$$

u kome je domet d izražen u kilometrima, a visina na kojoj je postavljena predajna antena - h, u metrima. U tabeli I je prikazana udaljenost horizonta za različite visine postavljanja antene.



Sl. 6. Ilustracija slabljenja u slobodnom prostoru.

TABELA I

Teorijski radio-horizont kada je antena postavljena na različitim visinama iznad površine Zemlje

Pozicija antene iznad nivoa zemlje [cm]	Domet [km]
10	1.13
30	1.95
60	2.77
100	3.57
200	5.05

U konkretnom eksperimentu koji smo izveli na otvorenom prostoru, pouzdana i uspešna veza je ostvarena između predajnika na brdu Čegar i prijemnika na lokaciji iznad Niške Banje, na rastojanju od oko 10 km, uz uslov da između njih postoji optička vidljivost (Slika 7). RSSI vrednost je tom prilikom iznosila oko -125.

Sl. 7. Deonica uspešnog prenosa na daljini većoj od 10 km kada postoji optička vidljivost.

VI. Zaključak

U radu su predstavljeni rezultati eksperimentalne procene dometa LoRa primopredajnika za IoT u urbanom i ruralnom okruženju. Dobijeni rezultati za urbanu sredinu pokazuju da je efekat senke dominantan uticaj u ukupnom slabljenju signala kada ne postoji optička vidljivost između predajnika i prijemnika, dok je uticaj fedinga tada znatno manji. U zatvorenom prostoru se usled ovoga veoma brzo smanjuje nivo signala sa povećanjem udaljenosti i broja prepreka, pa se može očekivati da će domet biti za nekoliko redova veličine manji od specifikacija za otvoreni prostor. Zbog toga je u zatvorenom prostoru pogodnije koristiti neke od tehnologija kratkog dometa, pri čemu treba obezbediti da predajnik i prijemnik budu fizički u međusobnoj poziciji kojom je omogućen uspešan prenos. Sa druge strane, u otvorenom prostoru se dobri rezultati u pogledu dometa mogu očekivati kada između predajnika i prijemnika postoji linija direktne optičke vidljivosti. U ovakvim slučajevima domet može biti i veći od 10 km, što je eksperimentalno potvrđeno na terenu. U slučaju da ne postoji optička vidljivost, ili da na liniji optičke vidljivosti postoje prepreke, domet će biti znatno manji, zavisno od specifičnosti određene situacije.

Zahvalnica

Projekat CANANDI finansira Fond za inovacionu delatnost iz budžeta Republike Srbije sa razdela Ministarstva prosvete, nauke i tehnološkog razvoja, a kroz Projekat za unapređenje konkurentnosti i zapošljavanja (sporazum o zajmu sa Svetskom bankom).

Literatura

[1] L. Farhan, R. Kharel "Internet of Things: Vision, Future Directions and Opportunities," In: Mukhopadhyay S., Jayasundera K., Postolache O. (eds) *Modern Sensing Technologies. Smart Sensors, Measurement and Instrumentation*, vol 29. Springer, Cham., 2019.

[2] E. Sisinni, A. Saifullah, S. Han, U. Jennehag and M. Gidlund, "Industrial Internet of Things: Challenges, Opportunities, and Directions," *IEEE Trans. Ind. Informat.*, vol. 14, no. 11, pp. 4724-4734, Nov. 2018.

[3] M. Sauter, *From GSM to LTE-Advanced Pro and 5G: An Introduction to Mobile Networks and Mobile Broadband*; Wiley: Hoboken, NJ, USA, 2017.

[4] A. Augustin, J. Yi, T.Clausen, W.M. Townsley, "A study of LoRa: Long range & Low Power Networks for the Internet of Things," *Sensors,* 16, 1466., 2016.

[5] I. Ahmed, S. Orfali, T. Khattab and A. Mohamed, "Characterization of the indoor-outdoor radio propagation channel at 2.4 GHz," *2011 IEEE GCC Conference and Exhibition (GCC)*, pp. 605-608, 2011.

[6] E. N. Sharma and G. C. Lall, "Enhancement and Character-ization of Indoor Propagation Models," *International Journal of Scientific & Engineering Research*, vol. 2, no. 11, pp. 1–7, Nov. 2011.

[7] E. Suikkanen, A. Tölli and M. Latva-aho, "Characterization of propagation in an outdoor-to-indoor scenario at 780 MHz," 21st *Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 70-74., 2010.

[8] P. M. Shankar, *Fading and Shadowing in Wireless Systems*, NewYork, NY, USA: Springer Science Business Media, 2012.

[9] B. Liu, B.P. Otis, S. Challa, P. Axon, C. Chou and S.K. Jha "The impact of fading and shadowing on the network performance of wireless sensor networks," *International Journal of Sensor Networks*, vol. 3, no. 4 pp. 211-223, 2008.

[10] S. Popa, N. Draghiciu and R. Reiz: "Fading Types in Wireless Communications Systems", *Journal of Electrical and Electronics Engineering*, vol. 1, pp. 232-237, 2008.

Abstract

In this paper we report practical results of signal level measurements during data transmission using a LoRa-based platform. Although the target applications for these platforms are primarily Internet of Things in urban environments, the paper also considers their use in rural surrounding. One of the ideas for an application is monitoring animals in environments that are located outside inhabited areas. Range of radio links in these environments can be significantly larger than in urban environments, depending on conditions of optical line-of-sight and obstacles located between the transmitter and the receiver. Practical results indicate that under favorable conditions acceptable transmission results can be achieved even at ranges beyond 10 km.

Range evaluation of LoRa IoT transceivers in urban and rural environment

Dejan N. Milić, Slavimir Stošović, Dejan Stevanović, Jelena Anastasov

On the impact of network load on CQI reporting and Link Adaptation in LTE systems

Igor A. Tomić, Milutin S. Davidović, Dejan D. Drajić, and Predrag Ivaniš

Abstract—**Data payload in mobile networks is persistently increasing, which puts a lot of pressure to Mobile Operators to seek for solutions that will deliver cheaper bit per second. Network performance modelling is important discipline in process of technology strategy definition, evaluation of different solutions/scenarios and finally design and planning of Long Term Evolution (LTE) systems. One of the key prerequisites for successful performance modelling process and prediction of user experience with growing network load is to understand impact of traffic increase on link adaptation. In this paper correlation between network load, measured as *Physical Resource Block* (PRB) utilization, and Link adaptation, measured with reported *Channel Quality Indicator* (CQI) is analyzed. Analysis is done based on performance measurements from mature commercial LTE network with several frequency layers deployed. Impact of network load increase on link adaptation performance was assessed. Performance of different frequency bands were observed separately, analysis was conducted for low band (bellow 1 GHz) and mid band (in range between 1 GHz and 6 GHz). Furthermore, analysis was segmented for different topologies in terms of network density – from urban to rural. Finally, impact of user mobility and spatial distribution of terminals, variations during working days and weekends were investigated.**

Index Terms—**LTE, Performance modelling, CQI, Network load, PRB utilization, Link adaptation, Spectral efficiency.**

I. Introduction

IN the last few years mobile operators were facing increase of data payload and growing network load as a major challenge in a mission to secure required user experience, that is mainly defined and measured as target downlink throughput or latency [1]. Fourth and fifth generation (4G and 5G) of mobile communication systems are based on *Orthogonal Frequency Division Multiple Access* (OFDMA) technique which enables flexible radio resource allocation [2], [3]. The most common answer from mobile operators has been the investment in spectrum expansions and deployment of additional frequency layers in carrier aggregation scenario.

Igor A. Tomić, University of Belgrade-School of Electrical Engineering, Bulevar kralja Aleksandra 73, 11000 Belgrade and Aspire Technology Unlimited – Belgrade Branch Office, Vladimira Popovića 6, Unit B10, 11070 Belgrade, Serbia (e-mail: igor.tomic@aspiretechnology.com).

Milutin S. Davidović, University of Belgrade-School of Electrical Engineering, Bulevar kralja Aleksandra 73, 11000 Belgrade and Aspire Technology Unlimited – Belgrade Branch Office, Vladimira Popovića 6, Unit B10, 11070 Belgrade, Serbia (e-mail: milutin.davidovic@aspiretechnology.com).

Dejan D. Drajić, University of Belgrade- School of Electrical Engineering, Bulevar kralja Aleksandra 73, 11000 Belgrade, Serbia (e-mail: ddrajic@etf.bg.ac.rs).

Predrag N. Ivaniš, University of Belgrade - School of Electrical Engineering,, Bulevar kralja Aleksandra 73, 11000 Belgrade, Serbia (e-mail: predrag.ivanis@etf.bg.ac.rs).

However, as spectrum assets are limited and scarce resources, price per Hz was increasing even dramatically on spectral auctions over last ten years, mainly due to growing demand which is lately present on all meridians around the globe [4]. In such circumstances, spectral efficiency is important more than ever [5], [6] and mobile operators are looking for spectral efficiency improvements to the largest possible extent.

The main driver for spectral efficiency is link adaptation, and its performance in OFDMA based systems is mainly driven by interference which is reflected in reported CQI based on *User Equipment* (UE) channel quality estimation and related measurements. With growing traffic, mainly for data services, but also speech and IoT [6], and higher network utilization observed, increase of intra-system interference is logically expected.

This paper investigates impact of growing network load on CQI reporting. In chapter II, theoretical considerations will be elaborated, while in chapter III impact of network load and correlation with CQI degradation will be investigated. Chapter IV will further analyze correlations for different network density, followed by deep dive to network performance signatures during working days and weekends in chapter V. Finally, paper will be concluded with overview of findings, proposal for future work and list of references.

II. Theoretical background

Link adaptation is the ability of radio communication systems to adapt the modulation scheme and the coding rate of the error correction according to the conditions that user expects and quality of the radio link. Algorithm will secure that a high-level efficient modulation scheme and a small amount of error correction is used when conditions of the radio link are good. Link adaptation happens in a downlink scheduler located in base station, which needs to know radio channel for each UE. The estimation of radio channel state and selection of optimal transmission scheme may be achieved through closed loop measurements, presented at Fig. 1, performed by UE on channels transmitted by base stations (*Cell reference Signal* – CRS in LTE) or reciprocity based on measurements conducted by base stations on signal transmitted by UE (*Sounding Reference Signal* – SRS), which is mainly applicable for Time Division Duplex (TDD) systems with good uplink-downlink channel reciprocity. In Frequency Division Duplex (FDD) systems closed loop channel estimation is used, where 3GPP standard requests each UE to perform *Channel State Information* (CSI) reporting that carries necessary information on a radio signal [8].

CSI reporting may be periodic and aperiodic. Periodic reporting is carried either by *Physical Uplink Control Channel* (PUCCH) or *Physical Uplink Shared Channel*

(PUSCH) when user data on *Dedicated Control Channel* (DCCH) and *Dedicated Traffic Channel* (DTCH) needs to be transmitted at the same time with L1/L2 control signaling. Periodicity is configured by a higher *Radio Resource Control* (RRC) layer. Aperiodic reporting is used mostly for *Random Access Channel* (RACH) procedure, where handovers or losses of synchronization are common reasons. Aperiodic reporting also happens when eNodeB schedules specific *Physical Downlink Control Channel* (PDCCH) *Downlink Control Indicator* (DCI) format 0 together with CQI request field set to 1, demanding uplink grant for UE data together with an aperiodic CQI report.

CSI consists of three major components: *Channel Quality Indicator* (CQI), *Precoding Matrix Index* (PMI) and *Rank Indicator* (RI). The focus of this paper is behavior of CQI, which implicitly indicates a suitable link adaptation and downlink transmission data rate, i.e. the highest *Modulation and Coding Scheme* (MCS) value at which the UE will be able to properly demodulate and decode the downlink data at target *Block Error Rate* (BLER) of 10%.

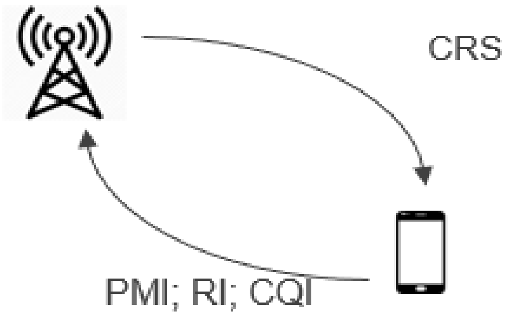


Fig. 1. Channel estimation based on Closed loop measurements, typical for LTE FDD systems

TABLE I
4-BIT CQI TABLE UP TO 64QAM MODULATION

CQI Index	Modulation	Code rate ×1024	efficiency
0	*Out of order*		
1	QPSK	78	0.1523
2	QPSK	120	0.2344
3	QPSK	193	0.3770
4	QPSK	308	0.6016
5	QPSK	449	0.8770
6	QPSK	602	1.1758
7	16QAM	378	1.4766
8	16QAM	490	1.9141
9	16QAM	616	2.4063
10	64QAM	466	2.7305
11	64QAM	567	3.3223
12	64QAM	666	3.9023
13	64QAM	772	4.5234
14	64QAM	873	5.1152
15	64QAM	948	5.5547

CQI calculation is not precisely defined by 3GPP standard and UE or chipset manufacturers have freedom to design algorithm and do implementation that will maximize accuracy and performance of link adaptation. However, calculations are typically based on measurements of *Signal-to-Interference plus Noise Ratio* (SINR), which are mapped to 16 discrete CQI values between 0 and 15, where index 15 indicates the best channel quality and index 1 indicates the poorest channel quality. Expected link adaptation - the mapping between CQI and modulation scheme and transport block size is defined by 3GPP [8], with two mapping tables defined for different UE capability. Table 1 is used for UEs supporting modulation up to 64QAM, while Table 2 is used for UEs supporting modulation up to 256QAM.

TABLE 2
4-BIT CQI TABLE UP TO 256QAM MODULATION

CQI Index	Modulation	Code rate ×1024	efficiency
0	*Out of order*		
1	QPSK	78	0.1523
2	QPSK	193	0.3770
3	QPSK	449	0.8770
4	16QAM	378	1.4766
5	16QAM	490	1.9141
6	16QAM	616	2.4063
7	64QAM	466	2.7305
8	64QAM	567	3.3223
9	64QAM	666	3.9023
10	64QAM	772	4.5234
11	64QAM	873	5.1152
12	256QAM	711	5.5547
13	256QAM	797	6.2266
14	256QAM	885	6.9141
15	256QAM	948	7.4063

CQI indicator is one of the main drivers for performance of LTE system, as it will determine link adaptation in terms of selected modulation, transmitted bits per symbol, coding and efficiency. Furthermore, there is a strong correlation observed between CQI and probability to have Rank higher than 1 and use more layers in *Multiple Input Multiple Output* (MIMO) system, which is often referred to as MIMO utilization [1]. Hence, lower CQI reported will cause additional negative impact on spectral efficiency, through worse spatial multiplexing performance.

III. CQI REPORTING AND NETWORK LOAD

Performance management data, including relevant counters and performance indicators, have been collected from commercial network that operates in two frequency bands: Band 12 - 700 MHz (operating DL frequency: 729 – 746 MHz) and Band 2 – PCS 1900 MHz (operating DL frequency: 1930–1990 MHz). Analysis is focused on average reported CQI and network load measured as average PRB utilization. Data was collected for period of ten days with cell resolution, where aggregation and averaging were performed afterwards on hourly level.

Fig. 2 presents correlation between average reported CQI and PRB utilization for two operating frequency bands, where first observation is a very good level of correlation between average CQI and network load, for both frequency bands, with correlation coefficient of 0.8173 for Band 2 and 0.9357 for Band 12. Network load is not evenly distributed, where for Band 12, average PRB utilization spans up to 45% during busy hour, while maximum measured PRB utilization for Band 2 is 20%.

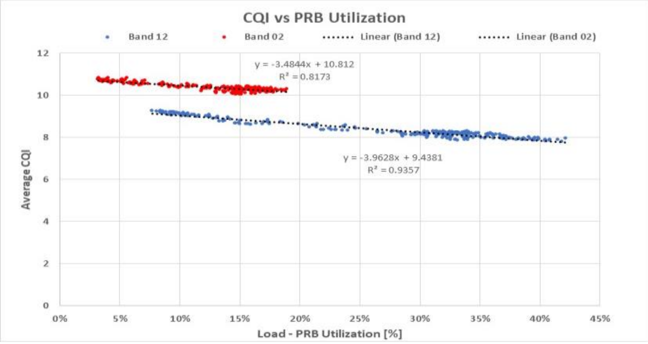


Fig. 2. Average reported CQI vs PRB utilization, LTE Bands 2 and 12

.

Some final remarks are that average CQI for same level of network load is higher on Band 2. This can be explained with lower operating frequency of Band 12(700 MHz vs 1900 MHz), which means that radio waves have better propagation in Band 12 and system is more sensitive on inter-cell interference. The difference between two curves in average reported CQI for same level of network load is quite significant – approximately 1.5.

From observation discussed in this chapter, it can be concluded that CQI reporting process in LTE network is driven mainly by inter-cell interference, with strong correlation with network load, and greater sensitivity for lower frequency bands.

IV. CQI Reporting and Network Density

Logical next step in analysis is to evaluate impact of network density in terms of number of deployed base stations and average Inter-Site Distance (ISD).

Network topology was analyzed with deep dive to geographical position of base stations, where sites were segmented to high and low network density area with following criteria:

- High network density area: if criteria more than three neighboring sites within radius of 1.2 km is met
- Low network density area: if criteria less than two neighboring sites within radius of 1.2 km is met

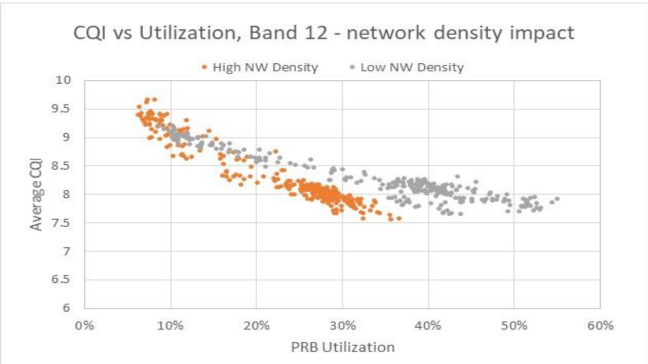


Fig. 3. CQI vs PRB Utilization, Network density impact, Band 12

Fig. 3 presents correlation between average CQI and PRB utilization, for two group of sites – belonging to high and low network density area, with operating frequency in Band 12. Similar analysis is done for Band 2, and results are presented on Fig. 4.

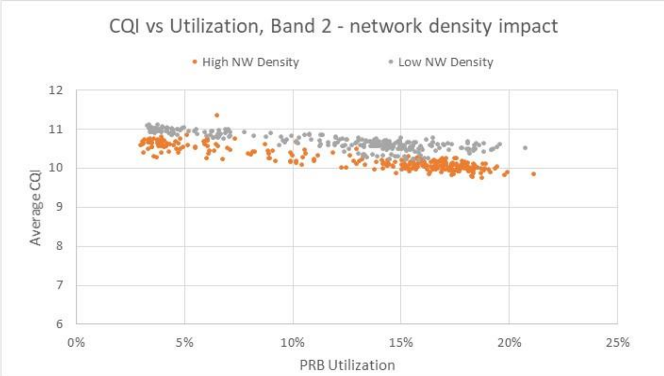


Fig. 4. CQI vs PRB Utilization, Network density impact, Band 2

The first observation is that base stations in high density areas are more affected with growing load, and lower reported CQI may be expected for same level of PRB utilization. Also, it is interesting to notice that impact of network density on correlation of interest is different for two observed frequency bands. While for Band 2 (Fig. 4) two curves are parallel, only shifted by approximately 0.5, in case of Band 12 (Fig. 3) steeper curve and higher relative drop may be expected in high density areas. For PRB utilization increase from 10% to 30%, in high density area average CQI reported value will drop from 9 to 8, while in low density area drop will be less severe, from 9 to 8.5.

V. CQI Reporting and User Distribution

After network load impact analysis, operating frequency band and network density on CQI reporting, another interesting aspect to consider is behavior of users in terms of their mobility and spatial distribution. One way to assess impact of mobility is to segment data for workdays – Monday to Friday, and weekends – Saturday to Sunday, as mobility patterns are clearly different.

As already noted in the Chapter I, good correlation is present in both Band 12 and Band 2 and it remained similar after the segmentation of the data sample to workdays and weekends.

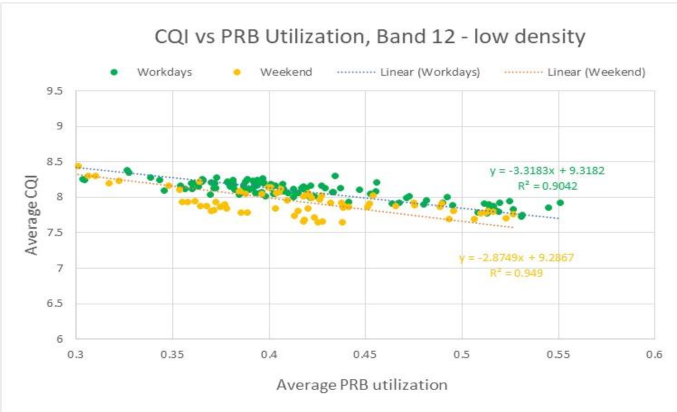


Fig. 5. CQI vs PRB Utilization, User Distribution Impact, Low density area, Band 12

However, there is a trend of having slightly lower CQI as well as the bigger deviation of reported CQI during the weekend.

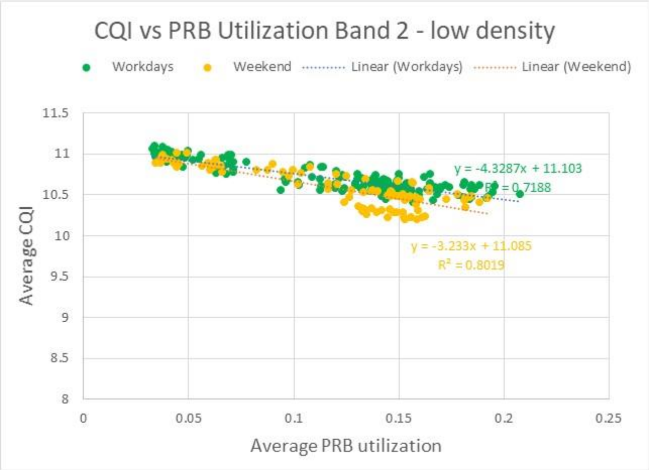


Fig. 6. CQI vs PRB Utilization, User Distribution Impact, Low density area, Band 2

This phenomenon is more noticeable in the low-density areas, especially in the range of higher PRB utilization in the observed sample (Fig. 5 and Fig. 6). The drop of 0.1-0.2 in average reported CQI during the weekend could be observed. In this case, another interesting phenomenon of having larger spam/deviation or even two trends of reported CQI for the same load circumstances could be observed. This could be due to different activity patterns of the users in some hours during the weekend in these areas, but further investigation (i.e. machine learning based) need to be conducted to drive some more specific conclusions.

On the other hand, looking at the high-density areas (Fig. 7 and Fig. 8), the impact of the day of the week on the CQI vs PRB utilization correlation is neglectable, most probably due to the reason that there is not much difference in user mobility and spatial distribution patterns in more dense areas of the network.

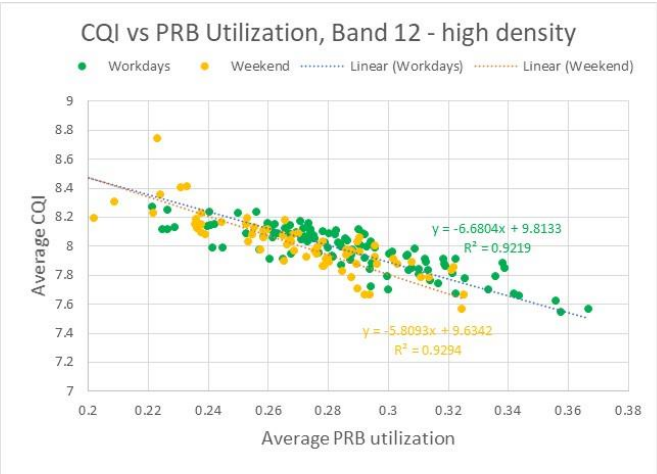


Fig. 7. CQI vs PRB Utilization, User Distribution Impact, High density area, Band 12

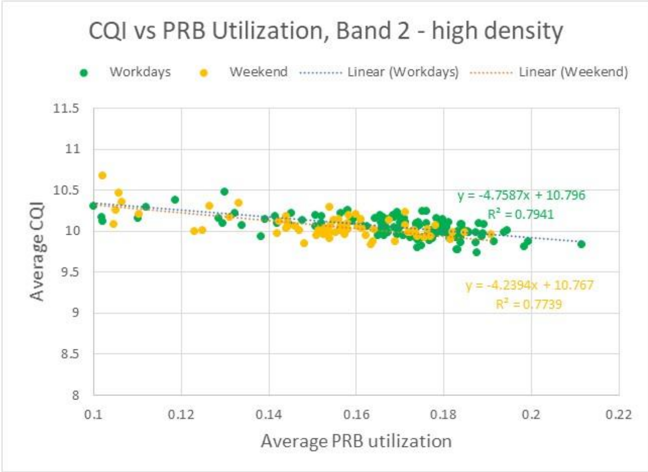


Fig. 8. CQI vs PRB Utilization, User Distribution Impact, High density area, Band 2

VI. Conclusion

Link adaptation is one of the main drivers for performance of OFDMA based mobile communication systems, such as LTE and NR. In this paper, the impact of network load on CQI reporting was analyzed and strong level of correlation was observed. Greater sensitivity for lower frequency bands was proven, which was in line with expectations, having in mind better propagation at lower frequencies causing more inter-cell interference. Furthermore, the areas with higher network density showed to be more affected with growing load and CQI degradation drop with steeper trend. Finally, impact of user mobility and traffic distribution is also evident, where different patterns were recognized during workdays and weekends.

Future work concerns deeper analysis of particular patterns through implementation of artificial intelligence and machine learning (AI/ML) techniques, and further correlation with spatial multiplexing performance in MIMO.

References

[1] I.Tomić, Đ.Lukić, M. Davidović, D. Drajić, and P. Ivaniš, “Statistical analysis of CQI reporting and MIMO utilization for downlink scheduling in live LTE mobile network”, *Telfor Journal 2020,* vol. 12, iss. 1, pp. 8-12

[2] C. Cox, *An Introduction to LTE, 2nd edition*, John Wiley & The Sons Ltd, Chichester 2014.

[3] N. Gospić, I. Tomić, D. Popović and D. Bogojević, *“Razvoj mobilnih telekomunikacija, od GSM do LTE,”* Univerzitet u Beogradu, Saobraćajni fakultet, Beograd 2010.

[4] Spectrum Navigator, Q1 2021 New insights and trends to watch, GSMA, May 2021, https://www.gsma.com/spectrum/wp-content/uploads/2021/03/Spectrum-Navigator-Q1-2021.pdf

[5] M.Davidović, S.Bjekovic and I.Tomic, “On the impact of network load on LTE network downlink performance”, *IcETRAN conference 2015*, Srebrno Jezero 2015.

[6] I.A.Tomić, M.S.Davidović and S.M.Bjeković, “On the downlink capacity of LTE cell,” *23rd Telecommunications Forum TELFOR 2015*, pp. 181-185, Belgrade 2015.

[7] M. Davidović, I. Tomić, D. Drajić, and Z. Čiča, “On the Impact of NB-IoT on LTE MBB Downlink Performance,” *TELFOR Journal*, Vol.11, No.1, pp. 20-24, Belgrade 2019.

[8] “*Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer Procedures,” 3GPP TS 36.213 15.7.0*, September 2019.

Design Problems in Implementation and Control of Malicious Drone Missions Jammers

Jovan Radivojević, Aleksandar Vujić, Mladen Mileusnić, Predrag Petrović, Aleksandar Lebl

***Abstract*— Three important problems related to the malicious drone missions jamming are analyzed in this paper. These problems are: 1) selection of drone signals which are going to be jammed and corresponding signal frequencies, 2) selection of optimum jamming strategies and definition of signal characteristics, and 3) design of reliable and user friendly system control mechanisms. The implemented solution allows jamming of a significant number of signals important for drone operation: video, telemetry and navigation signals. Classic sweeping is modified to multisweep jamming with continuous and discrete sub-bands to increase jamming efficiency. Device control is possible from the PC application software by the application of user friendly menus or from a specially designed module for remote control. The implemented jamming system is presented as well as the characteristics of the generated jamming signals.**

***Index Terms*—Malicious drone missions jamming; jamming scenario; remote control; user interface.**

I. Introduction

DRONES or unmanned aerial vehicle (UAV) are every day becoming more often implemented devices in replacing humans in dangerous and time wasting missions. As they come to the place of accident by air flight, sometimes they may save people and material resources just because they are faster than other technologies. But, drones are also used in various malicious missions. They may carry explosive to cause numerous victims and damages of many important objects and places such as airports, stadiums, residential areas, governmental facilities, commercial and industrial facilities, and so on. They may be applied for spy missions, smuggling illicit materials over borders or into and out of the prisons [1], [2]. Well prepared drone mission at one place may cause enormous economy loss and problems for the whole world [3] when just one concentrated attack disabled for a pretty long time 5% of world crude oil production. When considering data from Civil Aviation Directorate of the Republic of Serbia [4], more than 95% of drones in Serbia in 2017. were used illegally. All these illegal drones may be the reason of intentionally or unintentionally caused problem. Only these last two data from national and international literature and the short survey of places suitable for drone illegal activities are sufficient to approve the importance of drone effective jamming when their flight is not regularly registered.

Jovan Radivojević is with the Institute IRITEL, Batajnički put 23, 11080 Belgrade, Serbia (e-mail: jovan.radivojevic@iritel.com).
Aleksandar Vujić is with the Institute IRITEL, Batajnički put 23, 11080 Belgrade, Serbia (e-mail: aleksandar.vujic@iritel.com).
Mladen Mileusnić is with the Institute IRITEL, Batajnički put 23, 11080 Belgrade, Serbia (e-mail: mladenmi@iritel.com).
Predrag Petrović is with the Institute IRITEL, Batajnički put 23, 11080 Belgrade, Serbia (e-mail: presa@iritel.com).
Aleksandar Lebl is with the Institute IRITEL, Batajnički put 23, 11080 Belgrade, Serbia (e-mail: lebl@iritel.com).

II. The Problems to Be Solved

The problems which have to be solved to prevent illicit drone mission may be divided to three groups: 1) what drone signals have to be jammed; 2) what jamming strategies and jamming signal levels are the most effective for successful drone disabling; 3) what control scenario should be implemented to allow easy and reliable user handling the drone jammer.

There is a variety of drones available on the world market. All of them in principle operate using three main signal types: 1) communication signals interchange between the drone and its operator; 2) video and telemetry signals transmission to the operator; 3) satellite navigation signals reception (GPS or GLONASS). The frequencies for the third signal type are standardized, but the frequencies for the first two signal groups depend on the applied drone model. The more detailed analysis has shown that frequencies used for video and telemetry links at the greatest majority of drone types are 433 MHz, 868 MHz, 915 MHz, 1.2 GHz, 2.4 GHz and 5.8 GHz [2], [5] - [7]. The standardized frequencies for GPS and GLONASS systems are 1176 MHz, 1227 MHz and 1.57-1.62 GHz. It is hard to find that any drone jammer may effectively disrupt transmission of signals at each of the emphasized frequencies. Nearly all these frequencies are jammed in [6]; in other cases the dominant goal is to jam frequencies of navigation systems [7] or navigation system plus links on 2.4 GHz and 5.8 GHz [8]. The objective of the presented solution is to realize jamming of all emphasized frequencies for video and telemetry signals transmission in modern solutions and the frequencies for navigation. Such solution has the maximum flexibility and the highest possibilities.

Various jamming strategies are defined and developed for remote control improvised explosive devices (RCIED) activation jamming and also for radio and mobile systems surveillance and jamming. Two basic groups of jamming strategies are active and reactive jamming. The first of them supposes that jamming signal is continually generated. In the second case jamming signal is generated only when drone presence is detected. When considering signal characteristics, sweep and barrage jamming are most widely used as well as their variants and combinations. IRITEL has the great experience in both jamming development and implementation and theoretical analysis for all mentioned applications and signal generation strategies. References [9] - [18] represent the small part of IRITEL previous solutions relating to the topic of this paper. The goal of our analysis is to select a set of optimum jamming strategies and signal characteristics for effective jamming.

In the field of jammer control, the aim is to provide flexibility of operator handling. It means that the scenario of the first choice is to allow the parameters of the jamming signal to be easily and instantaneously changeable according to operator plans. In the more rigid situations only a set of jamming signal characteristics is defined and operator should make a decision between a priori specified options. This is the second device control procedure. The additional objective which results from these two scenarios is that reliability of control function practical implementation is significantly improved.

III. Design considerations

The frequencies which are jammed cover a wide frequency range 433-5800 MHz. It is important to notice that in the range 1620-2400 MHz there are no important frequencies to be jammed. Sweep and barrage jamming as two dominant jamming strategies have specific problems when jamming is applied to the wide bandwidth. When considering sweep jamming, the problem is to achieve very high sweep rate to reliably disrupt very short messages [13], [14]. In the case of pure barrage jamming implementation, the problem is necessity to have too high emission power which is split to the whole applied frequency band even where this is not necessary [15]. These two limiting factors are the reason to split the whole bandwidth into several sub-bands using separate jammer modules which may operate in parallel. We have decided to have five independent modules. Among these modules, the third module in the range 1164-1620 MHz covers the greatest number of frequencies of interest: all frequencies for navigation and one frequency for video and telemetry link. Only this module is sufficient to achieve performances comparable to the majority drone jammer solutions in the world.

Frequencies of the signals which have to be jammed are a priori mainly known. In such a case efficiency of sweep jamming may be improved. For example, it is not necessary to generate sweep signal in the whole third band 1164-1610 MHz, but only in some its parts where the frequencies of interest are located. It is possible to implement two sweep jamming strategies which decrease necessary sweep rate: 1) multisweep with continuous sub-bands (continuously sweeping over some parts of the total frequency band) and 2) multisweep with discrete sub-bands (sweeping only over discrete frequencies within frequency sub-bands). The second of these two variants is innovative in IRITEL jamming systems and not implemented in other solutions according to the authors' knowledge.

Flexibility of operator's jammer handling is achieved both by hardware and software methods. Hardware solutions allow variable gain adjust, precise jamming signal frequencies definition, more discrete frequencies selection for jamming, several (four) different jamming strategies implementation, precise definition of jamming and non-jamming time intervals and selection of one jamming scenario between several (eight) predefined programmed procedures when system is controlled by special developed remote controller. Dealing with software, user interface is intuitive, easily understandable and clear when control is devoted to a PC to support all wide hardware possibilities.

IV. The jammer implementation

IRITEL previous jammer solutions related to drones are presented in [19], [20]. In [19] it is explained how drones are used in friendly missions to improve RCIED activation jamming efficiency: if drone is carrying jamming device, the jamming signal may reach larger distance in comparison to the ground jamming. On the other hand, the main principles of malicious drone jamming system and the frequency spectrum of all implemented jamming strategies are presented in [20].

The block-scheme of the realized jamming system is now presented in the Fig. 1. The system includes the jamming device and separate module for remote control as also the application software for control and management at the distant PC computer.

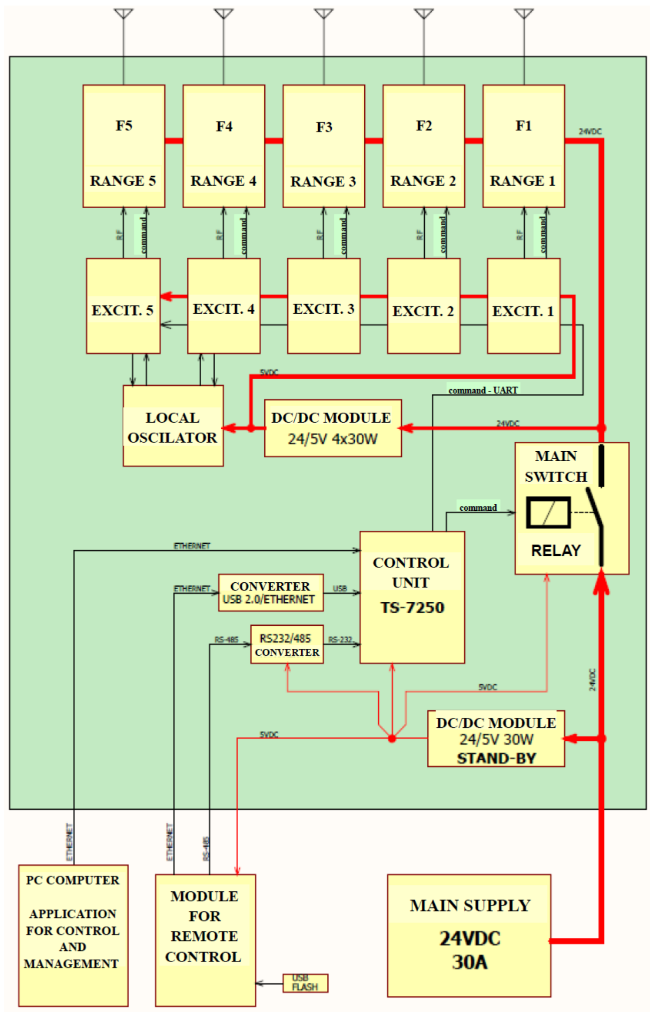


Fig. 1. Block-scheme of the malicious drone jamming system

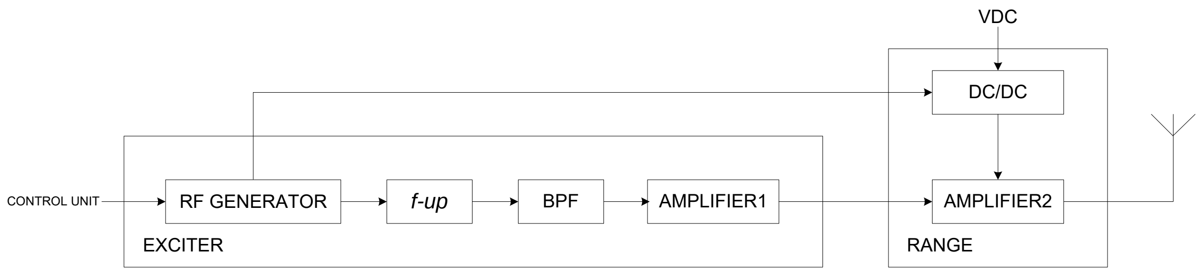


Fig. 2. Block-scheme of the EXCITER module and the output module RANGE

The jammer control and management is based on the implementation of processor board TS-7250. This module defines the functional characteristics of the EXCITER blocks for each of the five frequency bands. EXCITER blocks are based on the function of direct digital synthesis (DDS). After jamming signal generation and shaping in the exciter, the output signal level is achieved in the blocks designated by RANGE and further sent towards emission antennas. The signal frequencies in the last two ranges are too high for DDS realization and the function of these two segments is based on the lower frequency signal generation in the local oscillator. The signal in the local oscillator is used to shift the signal generated in DDS modules 4 and 5 to the desired frequency band by the implementation of mixers.

The control unit in the jammer exchanges information with the application software in distant PC and with the module for remote control using Ethernet and RS485 interface. Communication with the module for the remote control is realized over two converters: USB2.0/Ethernet and RS232/RS485.

The simplified, general block-scheme of the modules EXCITER and RANGE is presented in the Fig. 2. After generation of the jamming signal in RF GENERATOR, the signal frequency is up-converted (for the modules 4 and 5). Band pass filter (BPF) is used to finally achieve the desired jamming signal frequency. The obtained signal is then amplified by two amplifier stages, one in the block EXCITER and the other in the block RANGE to achieve the desired jamming signal level.

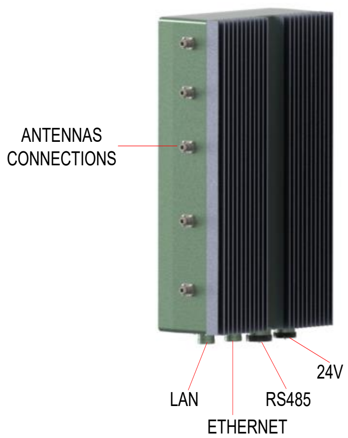


Fig. 3. Jammer device and its connections

The appearance of the jammer device is presented in the Fig. 3. Its dimensions are 640 mm x 320 mm x 155 mm. The jammer main connectors are designated in the figure as the connectors for five antennas, Ethernet and RS485 port for the connection to the remote control module and the LAN port to connect the remote PC. The last connector is for power supply 24 V. The measured device power supply current is 28.75 A from 24 V, thus obtaining total power consumption of 690 W when generating all five jamming signals. The system operating temperature range is from -20^{o}C till +55^{o}C. The time interval between the device switch-on till the start of jamming signal generation is 15s, while the time between the jamming signal characteristics definition change in a device menu till the moment of such signal generation start is 2s. Precision of jamming frequency definition is ±2kHz.

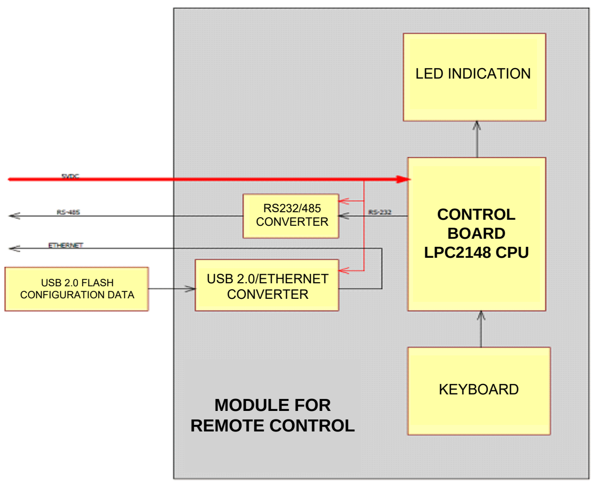


Fig. 4. Block-scheme of the module for remote control

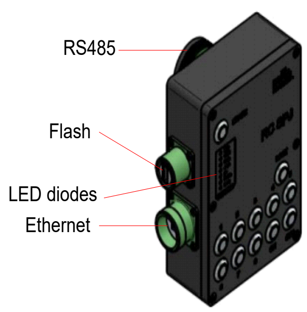


Fig. 5. Module for the remote control and its interfaces

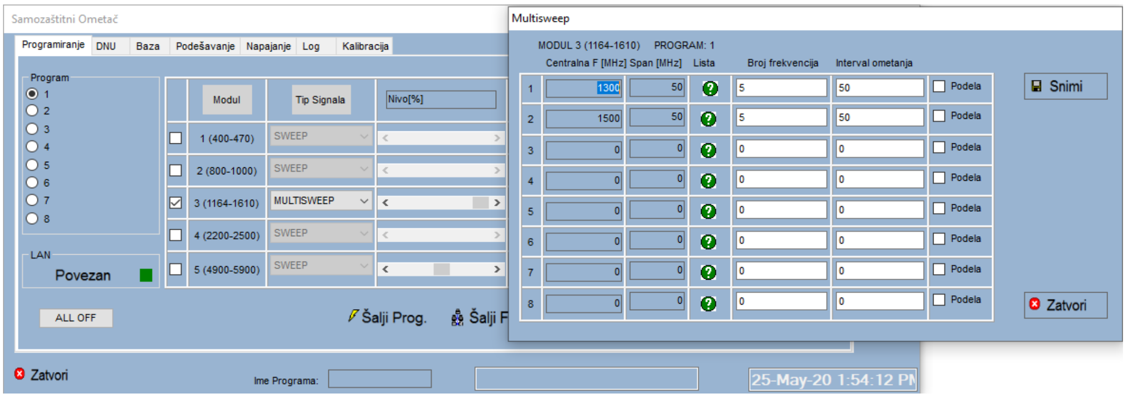


Fig. 6. User interface for jammer control when multisweep signal with discrete sub-bands is generated

The block-scheme of the module for the remote control is presented in the Fig. 4. The control board on the base of processor LPC2148 performs information transmission between the remote module and the device over RS485 and Ethernet interface. Dimensions of this module are 170 mm x 120 mm x 55 mm and its appearance is presented in the Fig. 5 with the designated main module parts.

The module for remote jammer control replaces control from the PC. Eight press buttons with the designation 1...8 for activating one of eight predefined scenarios are obvious at the nearer lateral side of this module. Ethernet and RS485 connector for connecting the remote control module and the jammer are the lower connector at the front side and the connector at the distant lateral side, respectively. The upper connector on the front side is intended for flash placement to read predefined jamming scenarios.

Fig. 6 presents user interface (menu) in the PC to generate multisweep signal with discrete sub-bands. In the presented example the signal is generated in the third band. The parameters of the signal may be very flexibly defined. There are two sub-bands (about 1300 MHz and about 1500 MHz and five frequencies in each sub-band). The span between two discrete frequencies is 50 MHz. Jamming for other frequency bands as well as for other jamming scenarios is also applicable.

V. Measurement Results

The measurement results for the third frequency band where there is a majority of jammed frequencies are presented in the Table I and in the Fig. 7. Table I presents the achieved maximum jamming signal power for all four jamming strategies. It is possible to define lower and even significantly lower emission signal level. This is especially important when considering GPS and GLONASS signal jamming. Navigation signals have very low level and it is enough to apply lower jamming signal levels. Too high jamming signal levels in this case could disrupt navigation signals at undesirably high distance outside the protected area [19], [21], where it is not necessary.

The Table II presents the maximum power of generated jamming signal for five signal sub-bands (F1 till F5) when sweep jamming strategy is applied. The similar results are obtained also for other three jamming strategies. F1 is the lowest frequency sub-band and F5 is the highest frequency sub-band. F3 covers the majority of applied frequencies for links of different type drones and these five sub-bands cover all drone frequencies mentioned in this paper.

TABLE I
THE POWER OF GENERATED JAMMING SIGNAL FOR VARIOUS JAMMING SCENARIOS

Jamming strategy	Generated signal power (dBm)	Generated signal power (W)
Sweep	46.4	43.7
Multisweep, continual sub-bands	46.4	43.7
Multisweep, discrete sub-bands	46.4	43.7
Barrage	46.2	41.7

TABLE II
THE MAXIMUM POWER OF GENERATED JAMMING SWEEP SIGNAL FOR VARIOUS FREQUENCY BANDS

Frequency band	Generated signal power (dBm)	Generated signal power (W)
F1	47.4	54.9
F2	46.2	41.7
F3	46.4	43.7
F4	44.6	28.8
F5	43	20

It is hardly possible to give exact presentation of jamming successfulness rate as a function of emitted jamming signal power, because it highly depends on the distance between the jammer and the drone. As a consequence, it is difficult to compare the jamming successfulness for various jammers from different suppliers. Two important factors which have great influence on the jamming successfulness, besides generated signal power in a jammer, are signal attenuation coefficient γ and existence/nonexistence of obstacles between the jammer and the drone. The influence of these two factors is even significantly higher than the influence of jamming power. For example, the signal power at a distance d from a jammer may be expressed as

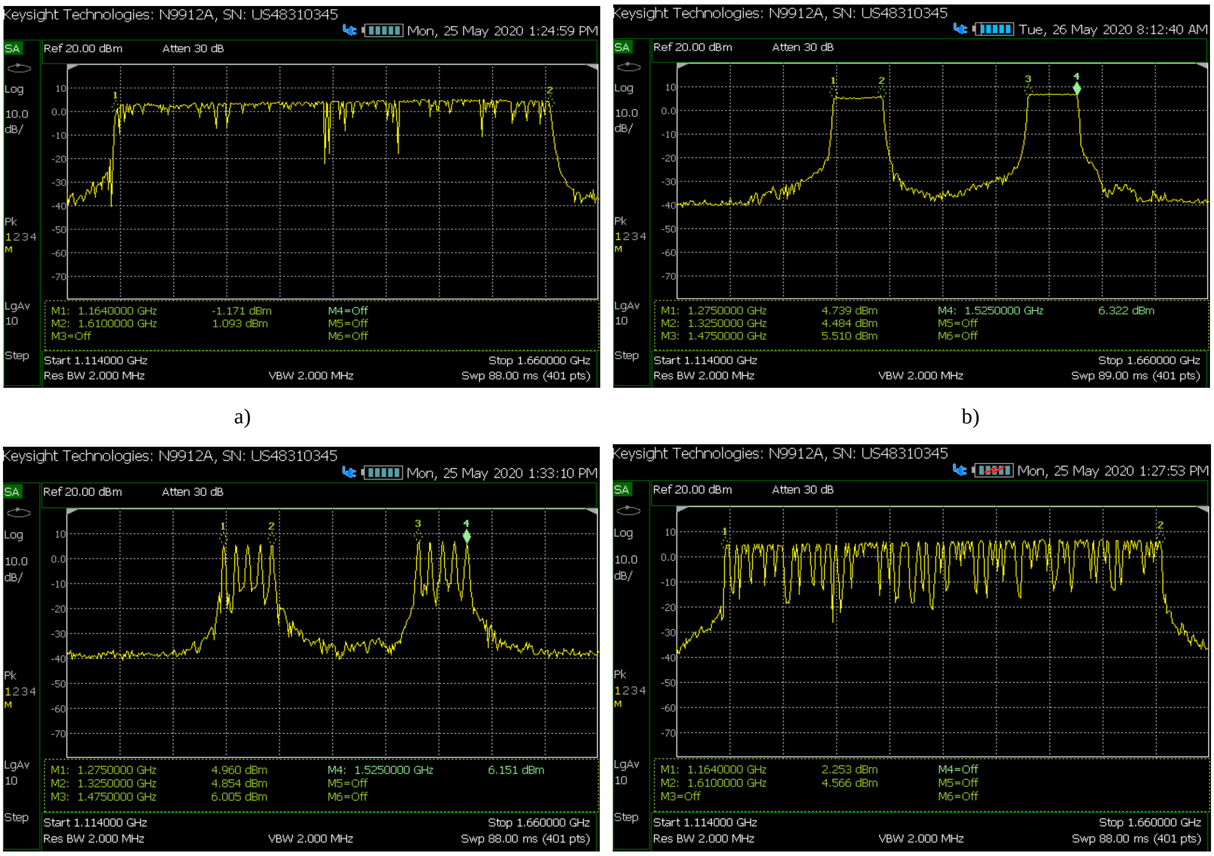
$$P(d) = a \cdot d^{-\gamma} \tag{1}$$

where a is the constant for adjusting values and dimensions of variables and γ has typical values between 2 and 5 [22]. It means that influence of propagation environment may cause signal attenuation in the ratio of even 8:1. Further, navigation signals should not be jammed by a too high level signal because navigation signals are also disrupted on other systems and devices where it is not desirable. As a brief illustration, according to [21] signal of the power 20mW (13dBm) is enough to block navigation signals in the range of 2km and this is significantly lower than the signal which our system may generate (Tables I and II). The jammer in [8] may cause successful jamming at 2km distance for directional jamming or 500m for omni-directional jamming. It means that our system may be effective at higher distance than solution [8].

Comparison in this case may be made when considering just signal levels at the place of signal reception. There is a number of papers whose authors include the authors of this paper where successful jamming rate (or bit error rate - BER) is determined as the level ratio RCIED activation signal/jamming signal [13]-[15], [18], [23]-[25]. Although the analysis in these papers is performed for RCIED jamming, the results may be also applied for drone jamming as the principles are the same. The results from these references prove that jamming successfulness depends on the applied jamming strategy (we have analysed the jamming strategies sweep, multisweep with continual sweeping and barrage in the mentioned papers) as well as on the jammed signal modulation type.

The frequency spectrum of the generated jamming signal is presented in the Fig. 7. for all four implemented jamming scenarios: sweep signal, multisweep signal with continuous and discrete sub-bands and barrage jamming. The signal spectrum is also related to the third frequency sub-band.

The applied jamming strategy mainly depends on the fact what data we have about the malicious drone and what drone signals we want to jam. In the case when a drone jammer is applied as a sole device, there is even no data whether a drone is present or not. It is necessary to jam a wide frequency spectrum in such a situation, perhaps by the jamming activation in all five sub-bands. The advantage of barrage jamming over sweep jamming when jamming is performed without any knowledge of drone presence is that all frequencies are jammed in the same time, but the necessary emitted power is higher than for sweep jamming [13], [14]. It means that, if jamming signal power is enough, jamming successful probability is 1 for barrage jamming and the sweep speed for sweep jamming must be determined in such a way that the probability of successful jamming is higher than some value, usually 0.95 [13], [14]. When the jammer is applied together with some system for drone detection and identification, we obtain some data about the malicious drone before jamming. Then it is possible to jam just on the frequencies where video, telemetry and drone communication signals are detected. It is possible to select to jam only one or two of these three signal groups, or only navigation signals. Thus emitted jamming signal power is significantly reduced.



a) b) c) d)

Fig. 7. Jamming signals in the frequency range 1164-1610MHz: a) sweep; b) multisweep with continual sweeping; c) multisweep with (five) discrete sweep steps; d) barrage

VI. CONCLUSION

The malicious drone missions jamming system presented in this paper completes IRITEL's palette of various implementation jammers [9] - [18]. With its wide flexibility in jamming strategies definition, different implementable jamming scenarios and suggestive handling menus, the presented jammer has comparable or even better characteristics than other internationally available solutions [7], [8], [26] - [30]. In the future the realized jamming system will be supplemented by the system for drones detection, identification, classification and localization to construct one comprehensive solution for the fight against malicious drones.

REFERENCES

[1] N. Eriksson, "Conceptual study of a future drone detection system Countering a threat posed by a disruptive technology," Master thesis in Product Development, Chalmers University of Technology, Goethenburg, Sweden, 2018.

[2] X. Shi, C. Yang, C. Liang, Z. Shi, and J. Chen, "Anti-Drone System with Multiple Surveillance Technologies: Architecture, Implementation, and Challenges," *IEEE Communications Magazine*, Vol. 56, Issue 4, April 2018., pp. 68-74., DOI: 10.1109/MCOM.2018.1700430.

[3] N. Razzouk, J. Blas, J. Thornhill, "Speed of Saudi Oil Recovery in Focus After Record Supply Loss," Bloomberg, 15. September 2019., https://www.bloomberg.com/news/articles/2019-09-15/saudis-race-to-restore-oil-output-after-crippling-aramco-attack.

[4] S. Milošević, "Nad Srbijom 4000 dronova, oni ilegalni biće obarani," („Over Serbia 4000 drones, the illegal ones will be destroyed"), *Novosti*, 15. October 2017., https://www.novosti.rs/vesti/naslovna/drustvo/aktuelno.290.html:690699-Nad-Srbijom-4000-dronova-oni-ilegalni-bice-obarani, in Serbian.

[5] A. L. Drozd, "Spectrum-secure Communications for Autonomous UAS/UAV Platforms," MILCOM 2015 - IEEE Military Communications Conference, Tampa, Florida, 26-28. October 2015.

[6] Drone Killer 6 – powerful UAV (GPS WIFI5GHz) Jammer – 120W, https://www.jammer-store.com/drone-killer-6.html.

[7] I. Pokrajac, N. Kozić, A. Čančarević, and R. Brusin, "Jamming of GNSS Signals," *Scientific Technical Review*, Vol. 68, No. 3, UDK: 621.396.96(047)=861, pp. 18-24, September 2018.

[8] Optix Anti-drone System, Product Information, Specification & Scope of Supply, Company Confidential.

[9] P. Petrović, N. Remenski, P. Jovanović, V. Tadić, B. Pavić, M. Mileusnić, B. Mišković, "WRJ 2004 wideband radio jammer against RCIEDs," tehničko rešenje – novi proizvod na projektu tehnološkog razvoja TR32051 pod nazivom "Razvoj i realizacija naredne generacije sistema, uređaja i softvera na bazi softverskog radija za radio i radarske mreže", 2011., http://www.iritel.com/images/pdf/wrj2004-e.pdf.

[10] "IRITEL High Frequency (HF) radio surveillance and jamming system," in the book M. Streetly, "*Jane's Radar And Electronic Warfare Systems*," IHS Global Limited, 2011.

[11] "IRITEL Very/Ultra High Frequency (V/UHF) radio surveillance and jamming system," in the book M. Streetly, "*Jane's Radar And Electronic Warfare Systems*," IHS Global Limited, 2011.

[12] M. Mileusnić, B. Pavić, V. Marinković-Nedelicki, P. Petrović, A. Lebl, "Development, realization and testing the variant of the device for VIP persons protection in the extended frequency range (20MHz-6GHz) with the reduced total power (300W)," tehničko rešenje – novi proizvod na projektu tehnološkog razvoja TR32051 "Razvoj i realizacija naredne generacije sistema, uređaja i softvera na bazi softverskog radija za radio i radarske mreže", 2017.

[13] M. Mileusnić, B. Pavić, V. Marinković-Nedelicki, P. Petrović, D. Mitić, A. Lebl, "Analysis of jamming successfulness against RCIED activation," 5th International Conference IcETRAN 2018, Palić, June 11-14, 2018., Proceedings of Papers, pp. 1206-1211, ISBN 978-86-7466-752-1, the best paper award in the Section of telecommunications.

[14] M. Mileusnić, B. Pavić, V. Marinković-Nedelicki, P. Petrović, D. Mitić, A. Lebl, "Analysis of Jamming Successfulness against RCIED Activation with the Emphasis on Sweep Jamming," *Facta Universitatis, Series: Electronics and Energetics*, Vol. 32, No. 2, June 2019., pp. 211-229., https://doi.org/10.2298/FUEE1902211M, extended and improved version of the paper awarded at 5th International Conference IcETRAN 2018, Palić, June 11-14. 2018.

[15] A. Lebl, M. Mileusnić, B. Pavić, V. Marinković-Nedelicki, P. Petrović, "Programmable Generator of Pseudo-White Noise for Jamming Applications," 27th Telecommunications Forum (TELFOR), Belgrade, November 26-27, 2019., Proceedings of Papers, pp. 1-4, ISBN 978-1-7281-4790-1, DOI: 10.1109/TELFOR48224.2019.8971203.

[16] M. Mileusnić, P. Petrović, A. Lebl, B. Pavić, "Comparison of RCIED Activation Responsive and Active Jamming Reliability," 6th International Conference IcETRAN 2019, Srebrno Jezero, June 3-5. 2019., Proceedings of Papers, pp. 988-993, ISBN 978-86-7466-785-9, the best paper in the Section of telecommunications.

[17] M. Mileusnić, P. Petrović, V. Kosjer, A. Lebl, B. Pavić, "Reliability Analysis of Different RCIED Activation Signal Responsive Jamming Techniques and Their Comparison to Active Jamming," *Facta Universitatis, Series: Electronics and Energetics*, Vol. 33, No. 3, September 2020., pp. 459-476., DOI: https://doi.org/10.2298/FUEE2003459M, extended and improved version of the paper awarded at 6th International Conference IcETRAN 2019, Srebrno jezero, June 3-5. 2019.

[18] A. Lebl, V. Kosjer, J. Radivojević, M. Mileusnić, "Performances of RCIED Activation Signal Multisweep Jamming," 7th International Conference IcETRAN 2020, Niš, September 28-29. 2020., Proceedings of Papers, pp. 740-745, ISBN 978-86-7466-852-8.

[19] V. Matić, V. Kosjer, A. Lebl, B. Pavić, J. Radivojević, "Methods for Drone Detection and Jamming," 10th International Conference on Information Society and Technology (ICIST), Kopaonik, March 8-11, 2020, In: Zdravković, M., Konjović, Z., Trajanović, M. (Eds.) ICIST 2020 Proceedings Vol.1, pp.16-21, 2020.

[20] J. Radivojević, A. Lebl, M. Mileusnić, A. Vujić, T. Šević, V. Joksimović, "Multichannel Radio-jammer Development Considerations for prevention of Illicit Drone Missions," 9th International Scientific Conference on Defensive Technologies OTEH 2020., Belgrade, October 15-16. 2020., pp. 270-275, ISBN 978-86-81123-83-6.

[21] D. A. M. da Silva, "GPS Jamming and Spoofing using Software Defined Radio," A Dissertation for the Degree of Master in Telecommunications and Computer Engineering, University Institute of Lisbon, 2017.

[22] J. Ebersprächer, H.–J. Vögel, C. Bettstetter, "GSM, Switching, Services and Protocols," (section 5), John Wiley & Sons, 1999.

[23] A. Lebl, B. Pavić, J. Radivojević, M. Mileusnić, P. Petrović, "Sweep Jamming Successfulness of DMPSK Modulated RCIED Activation Message," 28th Telecommunications Forum (TELFOR), Belgrade, November 24-25, 2020., Proceedings of Papers, pp. 21-24, ISBN 978-0-7381-4242-5, DOI: 10.1109/TELFOR51502.2020.9306587.

[24] M. Mileusnić, P. Petrović, B. Pavić, V. Marinković-Nedelicki, V. Matić, A. Lebl, "Jamming of MPSK Modulated Messages for RCIED Activation," 8th International Scientific Conference on Defensive Technologies OTEH 2018., Belgrade, October 11-12. 2018., pp. 380-385, ISBN 978-8681123-88-1.

[25] V. Marinković-Nedelicki, A. Lebl, M. Mileusnić, P. Petrović, B. Pavić, "BER Calculation for Sweep Jamming of MPSK Modulated RCIED Activation Message Signals," 18th International Symposium „INFOTEH Jahorina 2019", Jahorina, 20-22. March 2019., ISBN: 978-1-5386-7073-6, pp. 1-6, DOI: 10.1109/INFOTEH.2019.8717747.

[26] Optix: "Anti-Drone System Compact".

[27] https://www.perfectjammer.com/drone-signal-jammers.html.

[28] http://jammers4u.com/drones-jammer.

[29] https://www.thesignaljammer.com/blog/everything-you-need-to-know-about-drone-jammers/.

[30] Drone Killer 6 – powerful UAV (GPS WIFI5GHz) Jammer – 120W, https://www.jammer-store.com/drone-killer-6.html.

Performance Degradation of Coherent Direct Wideband Localization Due to Uncertainty in Receive Antenna Positions

Nenad Vukmirović, Miloš Janjić, and Miljko Erić

***Abstract*—In the context of passive coherent direct localization by a distributed receiving antenna array, we analyze how much the localization error increases due to non-ideal knowledge of receive antenna positions. We perform Monte-Carlo simulations with a wideband localization algorithm for a large distributed antenna array that surrounds the area were the transmitters are, and for an array of two pairs of antennas facing the area from a side. The former exhibits a very low increase in localization error, whereas the latter increases the error significantly, compared to the effect of the noise. We also derive approximate confidence intervals to confirm the validity of the drawn conclusions.**

***Index Terms*—Receive antenna position uncertainty; confidence intervals; coherent direct wideband position estimation; distributed antenna array; massive MIMO**

I. Introduction

IN this paper, we analyze a system that performs passive coherent direct wideband localization of a radio source (Tx) transmitting an arbitrarily wideband signal. The paper [1] explained the importance of wideband modeling, especially in newer generations of wireless systems. The (receive) antennas of the system are distributed in the area where the localization is performed. Therefore, we cannot assume planar wavefronts, but treat them as spherical, [2]. Each receive antenna is connected to an appropriate front-end, thus forming a single receive channel. Coherent localization requires that spatial coherence exists in the propagation medium and that the receive channels are time, frequency and phase synchronized, as described in [3], [4]. We assume that these conditions are satisfied. Even though the antennas are distributed, their front-ends could be held at the same place (collocated), such as in the example architecture in Fig. 1, which makes it easier to synchronize them. There are multiple sources of localization error in this scenario, such as the noise, interference, multipath propagation, synchronization errors and the uncertainty in the placement of the receive (Rx) antennas of the system, to mention a few.

The paper [3] showed that an error of about one thousandth of the carrier wavelength was achievable with coherent localization. A similar TDoA (Time Difference of Arrival) error (when converted to a length) was shown to be achievable in [5]. However, these results were obtained when the receive antenna positions were known exactly. The problem of accurate receive antenna placement is an important theoretical as well as practical problem, which the authors have encountered in a hardware implementation of a system for coherent localization, based on the methods in [3].

The impact of array element errors, either as (correlated) array shape distortions, independent errors of elements, or both, on the main beam direction, width, gain, as well as the sidelobe level was analyzed in [6]–[12]. The impact on direction of arrival estimation was analyzed in [13] and the impact on localization in [14]. Since the antennas in our paper are distributed, we model the element position errors as mutually independent and random. Also, since we are interested in localization of Txs inside the array aperture or close to it, the measure we use to quantify the impact of these errors is the localization accuracy.

In this paper, we are interested in the effects of the uncertainty in the Rx antenna positions on the accuracy of the mentioned type of localization. The authors of [15] proposed a method to estimate

N. V., M. J., and M. E. are with the University of Belgrade, School of Electrical Engineering, Serbia

N. V. and M. J. are also with the Innovation Center of School of Electrical Engineering, University of Belgrade, Serbia

M. E. is also with Vlatacom Institute, Belgrade, Serbia

This paper is a result of the research supported by the Serbian Ministry of Education, Science and Technological Development.

the position of a sensor in the array performing localization. The accuracy with which the antennas (the sensors) in the array are placed has to be greater in coherent radio localization, which this paper deals with, because, generally, the antennas have to be placed more accurately than the expected accuracy of the localization they are used for. Besides providing user location information for location-based services, the main purpose of coherent localization is to improve link performance in wireless systems, [16]. One such application of localization is in distributed beamforming. Distributed beamforming is robust with respect to the ambiguity problem, which exists in coherent localization (as explained on page 12 in [3]), so we do not focus on it in this paper. Many of the papers dealing with localization usually assume perfectly known Rx antenna positions, so it is important to analyze the effects of imperfect knowledge of Rx antenna positions. The results of this paper are important for selecting methods for measuring the geometrical relations (such as distances) between the antennas of the localization system, as well as for any applications that rely on accurate location estimation (with subwavelength accuracy) of radio transmitters.

II. Problem Formulation

Specifically, we want to quantify the increase in localization error due to the Rx antenna placement error compared to the noise-only scenario. To that end, we will use a signal model similar to that in [3], given by

$$u_m(t) = a_m \exp\left(-\mathrm{j}\omega_\mathrm{c}\left(t_0+\tau_m\right)\right) s\left(t-t_0-\tau_m\right)+\eta_m(t), \tag{1}$$

where $u_m(t)$ is the signal in Rx channel m, $m \in \{1,2,\ldots,M\}$, M is the number of the Rx antennas, a_m is a real-valued attenuation coefficient, $\omega_\mathrm{c} = 2\pi f_\mathrm{c}$ is the carrier frequency, t_0 is the unknown shift between the Tx and Rx time axis, $\tau_m = d_m/c$ is the propagation delay from the Tx, at an unknown position $\vec{r}$, to Rx antenna m, at $\vec{r}_m$, c is the speed of propagation, $d_m = \|\vec{r}-\vec{r}_m\|$, and $\eta_m(t)$ is the complex additive white Gaussian noise, AWGN. Figure 1 shows the system model and geometrical relations between the antennas. Note that, in coherent localization, the phase term contains the carrier phase only, and is modeled (by

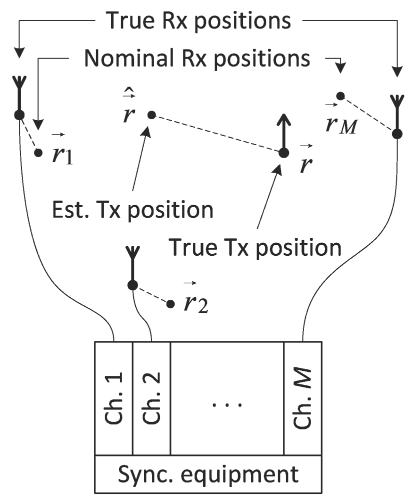


Fig. 1. The system model.

the exponential term) separately from the amplitude a_m. The term $s\left(t-t_0-\tau_m\right)$ models the envelope time delay of the transmitted signal/sequence $s(t)$ (wideband modeling). The signal-to-noise ratio, SNR, in channel m is $\mathrm{SNR}_m = \mathrm{SNR}_0 d_0^2/d_m^2$, where SNR_0 is used as a reference SNR at a distance of $d_0 = 1$ m. For convenience, let the unit of time be one sampling interval. This does *not* mean that the time variables/parameters are integers.

In each channel m, the samples available to a localization algorithm are $u_m(t)$, for $t \in \{0,1,\ldots,N-1\}$. The algorithm computes an estimate of the Tx position, $\widehat{\vec{r}}$, with an error $\Delta r = \|\widehat{\vec{r}}-\vec{r}\|$. The root-mean-squared localization error is

$$\mathrm{RMSE} = \sqrt{\mathrm{MSE}} = \sqrt{\mathrm{E}\,\Delta r^2}. \tag{2}$$

Since the Rx antenna positions, $\vec{r}_m$, are random variables, we can define the receive antenna RMSE, the RxRMSE, similarly. Since different methods can be used to position the Rx antennas, we use a generic model for the position errors. Let us assume that the error of each $\vec{r}_m$ along each of the axes (x and y; if antennas are distributed in 3D, then also z) is an independent 0-mean Gaussian random variable with the same variance σ_ax^2. Therefore RxRMSE is either $\sigma_\mathrm{ax}\sqrt{2}$ for 2D or $\sigma_\mathrm{ax}\sqrt{3}$ for 3D localization. The goal is to analyze how much the RMSE increases with RxRMSE.

III. Localization with Antenna Position Uncertainty: Error Analysis

We estimated the RMSE (which contained the effects of both the noise and the Rx antenna position uncertainty) through Monte-Carlo simulations for segments of random Gaussian transmitted sequences that were $N = 256$ samples long. The carrier frequency was 60 GHz and the bandwidth was 100 MHz. We used the SCM-MUSIC from [3] as a representative of coherent localization methods for unknown transmitted sequences. For a given set of Rx antenna positions, i.e. the array geometry, the true positions were randomly generated for each simulation run according to the given RxRMSE.

A. Approximate Confidence Intervals

We will characterize the quality of an estimate of the localization RMSE by an approximate confidence interval. To that end, we first note the distribution of the estimator $\widehat{\text{MSE}}$. We have that $\widehat{\text{MSE}}$ is the mean of the values Δr_k^2, $k \in \{1, 2, \ldots, K\}$, where K is the number of Monte-Carlo runs and Δr_k is the Euclidean distance between the estimated and the true location of the transmitter (i.e. the location error). The values Δr_k^2 are i.i.d. random variables, so, according to the Central limit theorem, the distribution of $\left(\widehat{\text{MSE}} - \mu\right)/\sigma$ is $\mathcal{N}(0,1)$, where μ and σ^2 are the mean and variance of $\widehat{\text{MSE}}$, respectively. Further, we have $\mu = \operatorname{E} \Delta r_k^2 = \text{MSE}$ (the true value) and $\sigma^2 = \operatorname{Var} \Delta r_k^2 / K$.

For a given level of confidence, p, define ε as $\mathcal{P}(|\xi| \leq \varepsilon) = p$, where $\xi \sim \mathcal{N}(0,1)$. Therefore,

$$\varepsilon = \sqrt{2}\, \operatorname{erfc}^{-1}(1-p). \tag{3}$$

Next, recall that

$$\Delta r_k^2 = \Delta x_k^2 + \Delta y_k^2 + \Delta z_k^2, \tag{4}$$

where Δx_k, Δy_k, and Δz_k are the individual errors along the coordinate system axes (for 2D localization $\Delta z_k = 0$). We assume that they are 0-mean Gaussian random variables with an unknown level of correlation. Thus, Δr_k^2 is expected to have a chi-square distribution with n degrees of freedom, $\chi^2(n)$, where we expect $n = 2$ in the case when the Tx is inside the aperture of the array for 2D localization, $n = 3$ inside the aperture for 3D localization, and $n = 1$ when the individual errors are highly correlated or when one of them is dominant over the other two (when the Tx is outside the array aperture). For a random variable $W_n \sim \chi^2(n)$, we rely on

$$\operatorname{Var} W_n = \frac{2}{n} \left(\operatorname{E} W_n\right)^2, \tag{5}$$

as a property of the chi-square distribution. Combining the previous properties, we obtain

$$\mathcal{P}\left(\left|\widehat{\text{MSE}} - \text{MSE}\right| \leq d\right) = p, \tag{6}$$

where

$$d = \varepsilon \sqrt{\frac{\operatorname{Var} \Delta r_k^2}{K}} = \text{MSE} \cdot \varepsilon \sqrt{\frac{2}{nK}}. \tag{7}$$

To approximate this, we use $n = 1$ as the worst case (the one that produces the widest confidence interval) and, since MSE is unknown, we use $\widehat{\text{MSE}}$ instead:

$$\widehat{d} = \widehat{\text{MSE}} \cdot \varepsilon \sqrt{\frac{2}{K}}. \tag{8}$$

Instead of using an absolute confidence interval $\left[\widehat{\text{MSE}} - \widehat{d}, \widehat{\text{MSE}} + \widehat{d}\right]$, we can use a relative one, $\left[\widehat{\text{MSE}}/\delta^2, \widehat{\text{MSE}} \cdot \delta^2\right]$, by defining δ as $\widehat{\text{MSE}}/\delta^2 = \widehat{\text{MSE}} - \widehat{d}$. This produces an approximate confidence interval for the $\widehat{\text{RMSE}} = \sqrt{\widehat{\text{MSE}}}$,

$$\left[\widehat{\text{RMSE}}/\delta, \widehat{\text{RMSE}} \cdot \delta\right], \tag{9}$$

where

$$\begin{aligned} \delta &= \sqrt{\frac{1}{1 - \varepsilon\sqrt{\frac{2}{K}}}} = \\ &= \left(1 - \frac{2}{\sqrt{K}} \operatorname{erfc}^{-1}(1-p)\right)^{-1/2}. \end{aligned} \tag{10}$$

Note that one convenient property of δ is that it does not depend on either the geometry, or the value of the MSE. It only depends on the number of simulation runs, K, and the chosen level of confidence, p. This formula is also useful for determining the number of simulation runs needed for a confidence interval of a given width.

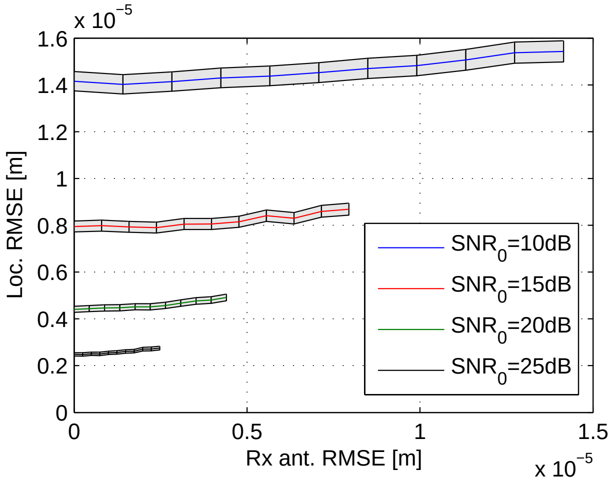


Fig. 2. The localization RMSE vs RxRMSE for different values of SNR_0 for the array geometry G_{18}.

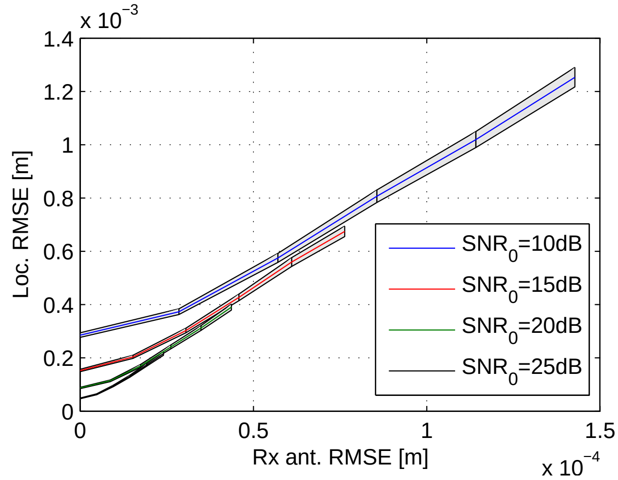


Fig. 3. The localization RMSE vs RxRMSE for different values of SNR_0 for the array geometry G_4.

B. Simulation Results

The results for different SNRs were generated for $K = 4096$ runs. The width of the confidence intervals (9), for $p = 0.99$, was then determined by $\delta = 1.03$, (10).

Figure 2 shows the results for a geometry G_{18}, used in [3], and based on [17], for localization in the horizontal plane 1.2 m below the array. The Tx was roughly below the center of G_{18}. The SNRs in the channels were grouped around the value 6 dB below SNR_0. The figure shows four curves for SNR_0 values of 10, 15, 20, and 25 dB. Note that they were evaluated only for RxRMSE below the RMSE of localization when there is no Rx antenna uncertainty, because it only makes sense that the accuracy of Rx antenna placement is greater than the accuracy of the localization method for the Tx. The curves show a very low increase in localization RMSE with an increase of RxRMSE. This can be explained by the fact that the number of Rx antennas is relatively large, they surround the area where the Tx is expected to be, and the placement errors are independent, so that all antennas would not move the main lobe of the localization algorithm in the same direction. Instead, those errors tend to partially cancel each other out, so the dispersion of the maximum of the main lobe is increased only slightly.

On the other hand, if some of the Rx antennas are close to each other (the antennas are grouped into subarrays), the number of antennas is small, or they do not surround the area where the Tx is, a much larger increase in RMSE is expected. Figure 3 confirms this and shows the results for a geometry G_4 with two subarrays of two antennas each, with their broadsides facing the area in front of the array. The (x, y) coordinates in [m] of the Rx antenna positions were $(0.0884, -0.0884)$, $(-0.0884, 0.0884)$, $(2.5316, -0.0884)$, and $(2.7084, 0.0884)$. The Tx was placed in front of the array at $(1.3, 1.5)$ in [m] and the distance to the Rx antennas was around 2 m. This means that the actual SNRs were 6 dB below SNR_0. To make the comparison fair, the curves were then evaluated for the same values of SNR_0 as for the G_{18} case. The increase in RMSE for the maximum mentioned RxRMSE was very large (around 9 times) so the curves are only shown for values below one half of that.

It would also be worthwhile to explore how the localization RMSE scales with the number of Rx antennas, m. However, to make the comparison for different values of m fair, for each m the array geometry should be optimized in some way. If the geometries were deterministic, it is unlikely that the geometry for m can be generated form the geometry for $m - 1$ by adding a single antenna without changing the positions of the others. If the geometries were random, the RMSE values should be averaged over different realizations of these geometries for all considered m. A uniform circular array could be considered for a fair comparison, but there are practical limitations that need to be considered as well, e.g., the antennas would probably be placed on walls or possibly on the ceiling of a room (in which the localization occurs), so this constrains the positions of the antennas to a rectangle/cuboid. However, since strict optimization of the Rx array geometry is outside the scope of the paper, this remains as an interesting topic for future research.

IV. Conclusion

We analyzed the impact of receive antenna position uncertainty on the accuracy of coherent direct wideband localization by a distributed receiving antenna array. Independent Gaussian errors in receive antenna positions were assumed. According to the simulation results, the impact of this uncertainty is small compared to the effect of the noise for G_{18}, which has a large number of antennas that encompass the area where the transmitters are. On the other hand, for arrays which have a small number of antennas, or have closely packed subarrays of antennas, especially when the transmitter is outside the aperture, such as G_4, the degradation of localization accuracy is rather large. The confidence intervals show that these effects are due to the Rx antenna positions uncertainty, and not merely due to the randomness in the simulations.

References

[1] B. Wang, F. Gao, S. Jin, H. Lin, G. Y. Li, S. Sun, and T. S. Rappaport, “Spatial-wideband effect in massive MIMO with application in mmWave systems,” *IEEE Communications Magazine*, vol. 56, no. 12, pp. 134–141, 2018.

[2] K. Han, Y. Liu, Z. Deng, L. Yin, and L. Shi, “Direct positioning method of mixed far-field and near-field based on 5G massive MIMO system,” *IEEE Access*, vol. 7, pp. 72 170–72 181, 2019.

[3] N. Vukmirović, M. Erić, M. Janjić, and P. M. Djurić, “Direct wideband coherent localization by distributed antenna arrays,” *Sensors*, vol. 19, no. 20, p. 4582, 2019.

[4] N. Vukmirović, M. Janjić, P. M. Djurić, and M. Erić, “Position estimation with a millimeter-wave massive MIMO system based on distributed steerable phased antenna arrays,” *EURASIP Journal on Advances in Signal Processing*, vol. 2018, no. 1, p. 33, 2018.

[5] M. Janjić, N. Vukmirović, and M. Erić, “TDOA, frequency and phase offsets estimation taking into account carrier phase of arrival,” *Radioengineering*, vol. 26, no. 4, pp. 1143–1150, 2017.

[6] X. T. Ye, W. T. Li, and W. H. Du, “An equivalent method of position error caused by the array antenna deformation,” *Progress In Electromagnetics Research Letters*, vol. 74, pp. 53–60, 2018.

[7] M. Mowlér, B. Lindmark, E. G. Larsson, and B. Ottersten, “Joint estimation of mutual coupling, element factor, and phase center in antenna arrays,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, no. 1, p. 30684, 2007.

[8] P. Marron, “Gain of phased array antennas under small random errors in element placement,” *M.D. thesis, Drexel University*, June 2014.

[9] T. Lindgren and J. Borg, “A measurement system for the position and phase errors of the elements in an antenna array subject to mutual coupling,” *International Journal of Antennas and Propagation*, vol. 2012, pp. 1–8, 2012.

[10] C. Wang, Y. Wang, X. Yang, W. Gao, C. Jiang, L. Wang, Y. Zhang, and M. Wang, “Effect of randomness in element position on performance of communication array antennas in internet of things,” *Wireless Communications and Mobile Computing*, vol. 2018, no. 2, pp. 1–8, 2018.

[11] M. Lanne, “Design aspects and pattern prediction for phased arrays with subarray position errors,” in *2010 IEEE International Symposium on Phased Array Systems and Technology*, 2010, pp. 440–446.

[12] H. S. C. Wang, “Performance of phased-array antennas with mechanical errors,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 2, pp. 535–545, 1992.

[13] R. Pöhlmann, E. Staudinger, S. Zhang, S. Caizzone, A. Dammann, and P. A. Hoeher, “In-field calibration of a multi-mode antenna for DoA estimation,” in *2021 15th European Conference on Antennas and Propagation (EuCAP)*, 2021, pp. 1–5.

[14] J. M. Fresno, G. Robles, B. G. Stewart, and J. M. Martínez-Tarifa, “The influence of antenna positioning errors on the radio-frequency localization of partial discharge sources,” in *3rd International Electronic Conference on Sensors and Applications*, 2016, pp. 1–7.

[15] J. C. Chen, R. E. Hudson, and K. Yao, “Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field,” *IEEE Transactions on Signal Processing*, vol. 50, no. 8, pp. 1843–1854, 2002.

[16] R. Di Taranto, S. Muppirisetty, R. Raulefs, D. Slock, T. Svensson, and H. Wymeersch, “Location-aware communications for 5G networks: How location information can improve scalability, latency, and robustness of 5G,” *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 102–112, 2014.

[17] “BK connect acoustic camera consisting of acoustic camera type 9712-W-FEN and BK connect array analysis type 8430,” https://www.bksv.com/-/media/literature/Product-Data/bp2534.ashx.

Coherent Method for Radio-Frequency Measurement of Distance between Antennas

Nenad Vukmirović, Miloš Janjić, Nikola Basta, and Miljko Erić

***Abstract*—The paper proposes a coherent method for radio-frequency measurement of the effective distance between two antennas. A transmitter sends a known waveform to a receiver, which processes the received signal to estimate its delay. The two transceivers are mutually synchronized, but different sources of delays/phase shifts still remain. A calibration step enables the system to estimate the total delay including the delays in the antennas. Differential measurements with three antennas enable us to estimate the delays in the antennas, which can be used as correction factors in measurements of the effective distance without the nuisance delays. The results of experiments performed on a prototype system, built of off-the-shelf equipment, show the consistency and the variance of the estimates. This method could be used to measure the geometry of a distributed array during its deployment, for the purposes of localization.**

***Index Terms*—Precise antenna positioning; coherent delay estimation; Universal Software Radio Peripheral; calibration; differential measurement**

I. Introduction

THIS paper proposes a radio-wave-based method for measuring the distance between two antennas and provides the results of experimental verification.

An approach for precise positioning of antennas in an ultra-wideband (UWB) system for indoor self-localization was discussed in [1]. It was found that a major source of localization error were the errors in the positions of the system's antennas. If the distance for each pair of antennas were measured accurately, the geometry of the distributed antenna array could be inferred and used for accurate localization. However, one should carefully consider which point of an antenna is the one that correctly represents it (the referent point). The best point is the one from which the radio wave propagates directly outward at the radio frequency which the antenna will be used for. This point is the antenna phase center. This complicates physical-distance measurements because the phase center may well be inside the body of the antenna (inaccessible for fine mechanical measurements) and it generally depends on the frequency as explained in [2], [3], [4]. These papers discuss measurement methods of the coordinates of an antenna's phase center and the complications that arise.

We propose a method based on RF (Radio Frequency) transmission and signal processing, which estimates the inter-antenna distance. This distance naturally corresponds to the phase centers and we call it the *effective distance*. Additionally, this allows us to circumvent the measurement of phase center positions and the complications associated with fine mechanical measurements of the distance between them. This represents an extension of the work in [5], which proposed a method for effective distance estimation based on group-delay and provided simulation results for different antenna orientations.

The paper [6] presented a method for coherent measurement of the distance between antennas (and the clock drift of their front-ends) up to an additive constant. We build on this idea by proposing a method that is also coherent, but is able to measure the absolute distance, thanks to a special calibration step and additional differential measurements that allow us to compensate for unwanted delays in the hardware, while having less stringent requirements on the synchronization of the front-ends. We use off-the-shelf equipment, including USRP (Universal Software Radio Peripheral) devices.

Coherent delay estimation, although potentially a lot more accurate than non-coherent methods, has an inherent ambiguity problem. Namely, the estimation error is a sum of an integral and fractional

N. V., M. J., N. B., and M. E. are with the University of Belgrade, School of Electrical Engineering, Serbia

N. V. and M. J. are also with the Innovation Center of School of Electrical Engineering, University of Belgrade, Serbia

M. E. is also with Vlatacom Institute, Belgrade, Serbia

This paper is a result of the research supported by the Serbian Ministry of Education, Science and Technological Development.

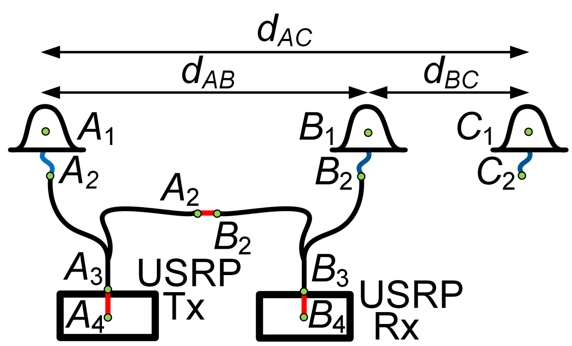


Fig. 1. The measurement setup with two USRP devices (Tx and Rx).

multiple of the carrier wavelength, the integral part representing the ambiguity. Even though the fractional error can be very small, applications that rely on accurate total distance measurements require the ambiguity to be resolved. The paper [7], although for a completely different setup (optical measurement), also has a method with the same error structure (see Sec. 2.3) and proposes a sophisticated method for resolving the ambiguity. In our paper we use a very coarse (and, therefore, simple) mechanical measurement of distance to resolve the ambiguity.

The most important application of the proposed method is measuring the (arbitrary) geometry of a distributed antenna system, for use in coherent localization [8], [9], which in turn can be used to improve the link performance of wireless systems, [10], by means of distributed beamforming or spatial multiplexing, to name a few.

II. Problem formulation

One USRP (Tx) transmits a predefined periodic waveform and another (Rx) receives the signal, which is then processed (see Fig. 1). One period of the waveform is determined by a vector of samples (a sequence), denoted by $\mathbf{s} = [s_0, s_1, \ldots, s_{N-1}]^\top$. Sequence $\mathbf{s}$ is a priori known at the Rx side and the processor estimates its delay with respect to the local Rx time axis.

The estimate is coherent, which means that the information is extracted not only from the signal envelope, but also from the carrier phase. This relies on the fact that when the delay of a signal in the RF domain increases by some $\Delta\tau$, the envelope is time-shifted by $-\Delta\tau$ and the carrier phase by $-\omega_c\Delta\tau$ (their shifts are coupled), where $\omega_c = 2\pi f_c$ is the carrier frequency.

Due to the restrictions of the off-the-shelf equipment used in the measurements, we have to consider different kinds of sources of error. A coherent algorithm requires the USRPs to be frequency synchronized. However, phase errors remain due to discrepancies in the lengths of the paths for the frequency reference (from a common source) in the two front-ends and the phase-locked loops (PLLs). There are also delays in the front-ends to (or from) the referent points in the chassis RF connectors (A_4A_3 and B_3B_4 in Fig. 1), in the signal cables between these connectors and the antenna connectors (A_3A_2 and B_2B_3), and between these and the referent points in the antennas (A_2A_1 and B_1B_2). Points A_1 and B_1 are not necessarily the phase centers of the antennas (in many antennas phase centers are not uniquely defined), but are chosen as geometrical reference points of antennas A and B. The algorithm considers that an RF wave propagates in free space between these points.

There is also a mismatch between the Tx and Rx local time axes, denoted by t_0. We wish to estimate the effective distance between antennas A and B, denoted by $d_{AB} = d\,(A_1B_1)$. We convert between delays and effective distances as $d_{AB} = c\tau_{AB}$, where c is the *free-space* propagation velocity of an RF wave at the given carrier frequency (by convention).

A phase drift exists, but its mean value is 0 and its variance is negligible because the propagation distances in the experiment are relatively short. There is also the clock drift, $\tau_{\text{clk}}(t)$, which can accumulate over time (its mean is not 0) because of independent clock generators in USRPs, but it only impacts the signal envelope. There can also be frequency-selective attenuation (e.g. due to anti-aliasing filters, coaxial cables, A/D or D/A converters...), which we separate into two factors – a delay (phase and time delay) and a waveform distortion factor (modeled by an unknown impulse response $h(t)$). Accounting for all these effects, we can write the model of the received signal, $u(t)$, as

$$\begin{aligned} u(t) =& s_r(t) + \eta(t) \\ s_r(t) =& a\mathrm{e}^{-\mathrm{j}(\omega_c\tau_a+\varphi)}\,(h*s)\,(t-\tau_a-\tau_{\text{clk}}(t)), \end{aligned} \tag{1}$$

where τ_a is the total constant envelope delay (which also includes the Tx t-axis shift t_0), $\omega_c\tau_a+\varphi$ is the total constant phase delay, $(h * s)(\cdot)$ is the convolution of the transmitted signal $s(t)$ with the impulse response $h(t)$ (modeling the effects in the entire path form the D/A converter in Tx to the A/D converter in Rx), $\tau_{\text{clk}}(t)$ is a process that slowly varies over time, a is a positive real-valued amplitude coefficient and $\eta(t)$ includes noise, multipath propagation and interference.

Note that we have normalized frequency values by the sampling frequency, $\widetilde{f}_{\text{s}}$, and time values by the sampling interval, $1/\widetilde{f}_{\text{s}}$, so that we can simply write $s(t) = s_t$, $t \in \{0, 1, \ldots, N-1\}$ (a convenient way to switch between the analog and the digital domain). This, however, does *not* restrict the argument for $s(\cdot)$ to a set of integers – $s(t)$ is still a continuous waveform. This also implies that c is normalized such that $c = \widetilde{c}/\widetilde{f}_{\text{s}}$ and has a unit [m/sample], where $\widetilde{c} = 3 \cdot 10^8$ m/s. If we need to explicitly write a parameter in physical units, we will use the symbol $\widetilde{}$ above it (e.g. $\widetilde{f}_{\text{c}} = 1.2$ GHz).

Since the estimation algorithm (described in Sec. III) is robust against the effects of $h(t)$, we will omit it. Also, since the algorithm is coherent, it perceives the phase delay φ as a time delay equal to $\varphi/\omega_{\text{c}}$. So, by substituting $\tau = \tau_{\text{a}} + \varphi/\omega_{\text{c}}$, we get a simplified model of the useful signal $s_{\text{r}}(t)$:

$$s_{\text{r}}(t) \approx a\mathrm{e}^{-\mathrm{j}\omega_{\text{c}}\tau} s(t - \tau - \tau_{\text{e}}(t)), \tag{2}$$

where τ is the total constant delay (for both the envelope and phase) and $\tau_{\text{e}}(t) = \tau_{\text{clk}}(t) - \varphi/\omega_{\text{c}}$ is the excess envelope-only delay (a slowly varying process).

III. The Measurement Method

A basic measurement consists of acquiring a signal segment of N samples, given by

$$\mathbf{u} = [u(0), u(1), \ldots, u(N-1)]^\top \tag{3}$$

and estimating the delay of the sequence $\mathbf{s}$ within that segment by the GCC-type (Generalized Cross-Correlation) delay-estimation algorithm given in [8],

$$\widehat{\tau} = \underset{\tau \in [0,N)}{\arg\max} \operatorname{Re}\left(\mathbf{u}^{\text{H}} \mathbf{s}_{-\tau}\right) \tag{4}$$

where $\mathbf{s}_{-\tau} = \mathrm{e}^{\mathrm{j}\omega_{\text{c}}\tau} \left[s(\tau), s(\tau+1), \ldots, s(\tau+N-1)\right]^\top$ is a vector of samples of the waveform $s(t)$ time-shifted by τ (or delayed by $-\tau$) *including* the coupled carrier phase shift, Re denotes the real part, $^\top$ matrix transpose, and $^{\text{H}}$ conjugate transpose. Since $s(t)$ is periodic with period N, a regular time shift is the same as a cyclic time shift, thus we can compute $\mathbf{s}_{-\tau}$ using DFT (Discrete Fourier Transform) as

$$\mathbf{s}_{-\tau} = \mathbf{F}^{\text{H}} \mathbf{D}_{-\tau} \mathbf{F} \mathbf{s}, \tag{5}$$

$$\mathbf{F} = 1/\sqrt{N} \exp(-\mathrm{j}2\pi/N \cdot \mathbf{k}\mathbf{n}^\top), \tag{6}$$

$$\mathbf{D}_{-\tau} = \exp\left(\mathrm{j}\omega_{\text{c}}\tau\right) \operatorname{Diag}\left\{\exp\left(\mathrm{j}2\pi/N \cdot \mathbf{k}\tau\right)\right\}, \tag{7}$$

$$\mathbf{n} = [0, 1, \ldots, N-1]^\top, \tag{8}$$

$$\mathbf{k} = [-N/2, -N/2+1, \ldots, N/2-1]^\top. \tag{9}$$

$\mathbf{F}$ is a modified DFT matrix such that $\mathbf{F}^{-1} = \mathbf{F}^{\text{H}}$, so a numerically more efficient form of (4) is

$$\widehat{\tau} = \underset{\tau \in [0,N)}{\arg\max} \operatorname{Re}\left(\mathbf{U}^{\text{H}} \mathbf{D}_{-\tau} \mathbf{S}\right), \tag{10}$$

where $\mathbf{U} = \mathbf{F}\mathbf{u}$ and $\mathbf{S} = \mathbf{F}\mathbf{s}$ are computed only once in preprocessing.

A. Type-I Measurement

In the *antenna configuration* in Fig. 1, in which the signal cables A_3A_2 and B_2B_3 are connected to the antennas, a basic measurement provides an estimate

$$\widehat{\tau}_{\text{ant}} = \tau_{\text{ses}} + \tau_A + \tau_{AB} + \tau_B + \varepsilon_{\text{ant}}, \tag{11}$$

where τ_A is the delay in antenna A (from A_2 to A_1), τ_B in antenna B (i.e. B_1B_2), τ_{ses} is the combined effect of the signal cables and front-ends, and ε_{ant} is the error.

The delay τ_{ses} is preserved within one driver session with the USRP devices (if we neglect the effects of the drift $\tau_{\text{e}}(t)$), but it takes a new value (changes unpredictably) each time a new session starts. Therefore, we can perform another basic measurement within *the same* driver session, but in the *guided configuration* (see Fig. 1), in which the signal cables are connected to each other by a short connector of effective length $L = c\tau_L$ (between A_2 and B_2). We get an estimate

$$\widehat{\tau}_{\text{g}} = \tau_{\text{ses}} + \tau_L + \varepsilon_{\text{g}}, \tag{12}$$

where ε_{g} is the error. From (11) and (12) we get

$$\tau_{AB} = \widehat{\tau}' - \tau_A - \tau_B + \tau_L + \varepsilon', \tag{13}$$

where $\widehat{\tau}' = \widehat{\tau}_{\text{ant}} - \widehat{\tau}_{\text{g}}$, $\varepsilon' = \varepsilon_{\text{g}} - \varepsilon_{\text{ant}}$ is the combined error. Note that we can perform multiple basic measurements in each of the two configurations to get more accurate estimates of $\widehat{\tau}_{\text{ant}}$ and $\widehat{\tau}_{\text{g}}$, thanks to averaging. Also, since the drift in $\tau_{\text{e}}(t)$ accumulates over time, by making the session shorter, i.e. if we change the configurations (disconnect and reconnect the cables) more quickly, the negative effect of $\tau_{\text{e}}(t)$ is expected to be smaller. Taking multiple measurements can also be useful as a way to diagnose this negative effect, especially if they are spread out across the interval of the session, because this effect is expected to grow over time.

We call the procedure of acquiring one or more $\widehat{\tau}_{\text{ant}}$ and one or more $\widehat{\tau}_{\text{g}}$ estimates, all within the same driver session, a type-I measurement.

B. *Type-II Measurement*

We can obtain τ_L in (13) in a fine measurement by some external means, but the terms τ_A and τ_B remain unsolved. We propose to use antenna C, with the referent point C_1, such that A_1, B_1, and C_1 are colinear as in Fig. 1. In that setup, we would perform a type-I measurement with antennas A and B and another with B and C. We would then remove the antenna B, so that it does not block the propagation from A to C and perform a type-I measurement with A and C. As a result, we get

$$\tau_{AB} = \widehat{\tau}'_1 - \tau_A - \tau_B + \tau_L + \varepsilon'_1, \tag{14}$$
$$\tau_{BC} = \widehat{\tau}'_2 - \tau_B - \tau_C + \tau_L + \varepsilon'_2, \tag{15}$$
$$\tau_{AC} = \widehat{\tau}'_3 - \tau_A - \tau_C + \tau_L + \varepsilon'_3. \tag{16}$$

Relying on $\tau_{AB} + \tau_{BC} = \tau_{AC}$, we get an estimate of τ_B as

$$\widehat{\tau}_B = \tau_B + \varepsilon'' = (\widehat{\tau}'_1 + \widehat{\tau}'_2 - \widehat{\tau}'_3 + \tau_L)/2. \tag{17}$$

where ε'' is the combined error. We call this procedure of obtaining $\widehat{\tau}_B$ a type-II measurement.

Note that $\widehat{\tau}_B$ is independent of the distances between the antennas and the delays in antennas A and C, for fixed signal-to-noise ratios, as long as the antennas are in the far-field regions of one another. Additionally, a deviation of B_1 from the line A_1C_1 influences ε'' as a second-order infinitesimal, so this deviation can be neglected.

Type-II measurements can be performed for each of the antennas in an RF anechoic chamber before they are deployed. Once deployed, the estimates $\widehat{\tau}_A$, $\widehat{\tau}_B$, ... can be input into (13) as correction factors, providing

$$\widehat{\tau}_{AB} = \widehat{\tau}' - \widehat{\tau}_A - \widehat{\tau}_B + \widehat{\tau}_L. \tag{18}$$

This way we get an estimate of the effective electrical distance between A and B, $\widehat{d}_{AB} = c\widehat{\tau}_{AB}$, or any other pair of antennas we use to form a distributed antenna array. As a result, the antennas can be placed arbitrarily and then the geometry of the array can be *measured*, by multiple type-I measurements (with appropriate correction factors).

C. *Integral and fractional errors*

In each of these measurements, the error (such as ε' in (13)) has two components – the first is an integer multiple of the carrier cycle, $1/f_{\text{c}}$, for delays, or the carrier wavelength, $\lambda_{\text{c}} = c/f_{\text{c}}$, for lengths, and the second component is the fractional part (the remainder, which is usually a few orders of magnitude smaller than λ_{c}):

$$\begin{aligned} \varepsilon' &= m/f_{\text{c}} + \varepsilon'_{\text{frac}}, \text{ for delays} \\ c\varepsilon' &= m\lambda_{\text{c}} + c\varepsilon'_{\text{frac}}, \text{ for distances}, \end{aligned} \tag{19}$$

for some $m \in \mathbb{Z}$. The presence of $m\lambda_{\text{c}}$ is known as the (integer wavelength) ambiguity problem and it is characteristic of coherent delay algorithms [11], [12]. This effect occurs because an RF signal and its replica delayed by $1/f_{\text{c}}$ closely resemble each other. The term $\tau_{\text{e}}(t)$ increases the error, but mostly its integral part (the ambiguity). Even though the integral part of the error is usually much larger, it is not as important as the $\varepsilon'_{\text{frac}}$ part. For example, the performance of distributed beamforming is deteriorated mostly by $\varepsilon'_{\text{frac}}$ and very little by the ambiguity.

We propose to solve the ambiguity problem by taking a rough estimate of the distance, d_{AB}, by some external means (mechanical means, laser range finder, or even an RF measurement but at a different carrier frequency). That method only has to be accurate enough to correctly resolve m (e.g., at $\widetilde{f}_{\text{c}} = 1$ GHz, $\lambda_{\text{c}} = 30$ cm, an accuracy of a few centimeters would suffice).

Finally, the effective lengths/distances of the guided parts of the system can depend on $\widetilde{f}_{\text{c}}$, so care should be taken to measure them at different

Fig. 2. The experiment environment.

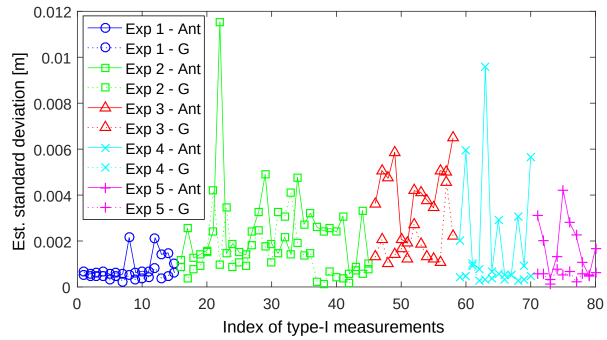


Fig. 3. Estimated standard deviations of lengths obtained within different type-I measurements.

frequencies and later use the correction factors at the correct frequency.

IV. Results of Measurements

We performed experiments with n210 USRP devices with omnidirectional antennas (connected by 25 m long coaxial cables) in an outdoor urban environment in front of the Innovation Center of the School of Electrical Engineering in Belgrade (Fig. 2). Most of the type-I measurements consisted of 8 and 4 basic measurements in the antenna and guided configuration, respectively. To reduce the bias in the estimates, we did not use an attenuator to connect A_2 and B_2. Thus, the Rx signal level was significantly lower in the antenna configuration than in the guided one. So, we used a greater number of measurements (8) in the antenna configuration to compensate for higher error variance. We performed multiple different experiments as a proof-of-concept.

In experiment (exp.) 1, we performed a type-II measurement with antennas A, B, and C in a line in that order, at $\widetilde{f}_{\mathrm{c}} = 990$ MHz. There were 5 type-I measurements for each pair of antennas (AB, BC, and AC). Fig. 3 shows standard deviations, σ_{A} and σ_{G}, of the effective length estimates in the antenna and guided configuration, respectively, within each type-I measurement. In this section, we concentrate on the std. deviations of the fractional part $c\varepsilon'_{\mathrm{frac}}$ from (19) (throughout these experiments, the std. deviations of the ambiguity m in (19) were grouped around 1 and most of them were in the range between 0 to 3). In most cases, σ_{A} was greater than σ_{G}, as expected, and both were less than 7 mm. The resulting 3 groups of estimates (one for each antenna pair) had standard deviations 1.5 mm, 3.5 mm, and 2.9 mm, as shown in Table I. The final (type-II) estimate of exp. 1 (the effective length of antenna B) was $L_B = 54.94$ cm, according to (17). (This was expected because the physical length of B_1B_2 was around 40 cm, with a roughly 1.5 times lower propagation velocity.) Note that this was a two-stage averaging, which made the total number of basic measurements 120 (antenna) and 60 (guided) for this single type-II estimate.

TABLE I
Parameters of experiments and standard deviations (σ).

Exp.	Type	Ant. pair	No. of meas.	σ [mm]	$\sigma/\lambda_{\mathrm{c}}$
1	II	AB	5	1.5	0.005
		BC	5	3.5	0.012
		AC	5	2.9	0.0096
2	II	AB	10	7.7	0.025
		BC	10	2.4	0.008
		AC	10	7.8	0.026
3	II	AB	5	9.3	0.034
		BC	4	3.9	0.014
		AC	4	3	0.011
4	II	BA	4	7.6	0.025
		AC	4	2.6	0.0086
		BC	4	11.3	0.037
5	I	AB	10	5.6	0.018

To reduce the number of times the cables had to be disconnected/reconnected to a half (less wear-and-tear of connectors), we started one type-I measurement in the antenna configuration and ended it in the guided one, and then reversed the order in the next measurement. We continued alternating this order throughout the campaign. Another advantage of this was to partially compensate for the bias created by a clock drift growing in one direction over the course of a driver session. The expected effect of this was an increase in σ of type-I estimates for each pair of antennas and a decrease in the bias in the final estimate of the pair.

To test the repeatability of measurement, exp. 2 was with the same antennas and $\widetilde{f}_c$ as in exp. 1, but in different conditions – different weather conditions (day) and inter-antenna distances. There were 10 type-I measurements for each antenna pair and the σ values of each pair (as well as the ones for the rest of the experiments in this section) are given in Table I. For individual type-I σ values, see Fig. 3. The final estimate was $L_B = 55.08$ cm.

To test the consistency across the frequencies, exp. 3 was carried out with the same setup as exp. 2, but at $\widetilde{f}_c = 1.1$ GHz (in this experiment only). Note that the effective length B_1B_2 (L_B) was not expected to be identical at these two frequencies (even though the physical one was). The final estimate was $L_B = 56.04$ cm.

In exp. 4 the antennas were placed as B-A-C (antenna A in the middle), with the goal of estimating L_A. The final estimate was $L_A = 66.68$ cm.

Exp. 5 consisted of 10 type-I measurements with antennas A and B, but with correction factors from exps. 2 and 4 (using (18)). The aim was to estimate the effective distance between the referent points of these antennas (i.e., in the air only), L_{AB}. According to a mechanical measurement, the distance was 3.45 m. The resulting individual σ values are given in Fig. 3 and the σ value of the final result in Table I. The final estimate was $d_{AB} = 3.32$ m.

In future work, one might consider decreasing the difference in the Rx signal levels in the antenna and guided configuration, in order to reduce the influence of thermal and quantization noise, but without inducing a bias in the delay estimation. Furthermore, compensation of the remaining bias in the estimates is also of interest. Additionally, (dis)connecting cables manually is impractical for deployments of distributed antenna arrays, so the switching (between configurations) may be designed so it would be electronically controlled and automatic, thus decreasing the probability of human error in the experiment.

V. Conclusion

In this paper we proposed a method for measuring electronically the effective distance between a pair of antennas. A prototype was built of off-the-shelf equipment. The results of the field tests showed the potential of the presented method, as well as the aspects that require improvement in order to use it in deploying distributed antenna arrays with subwavelength accuracy. Such a method can have a great impact on future wireless systems as an enabler for coherent localization.

References

[1] M. Erić, R. Zetik, and D. Vučić, "An approach for determination of antenna positions in distributed antenna system used for UWB indoor self-localization: Experimental results," in *2013 21st Telecommunications Forum Telfor (TELFOR)*, 2013, pp. 204–207.

[2] H. Yuan, H. Liu, J. Jia, T. Su, and C. Zeng, "A novel algorithm to compute 3d phase center," *Electromagnetics*, vol. 40, no. 7, pp. 486–499, 2020.

[3] P. N. Betjes, "An algorithm for automated phase center determination and its implementation," in *AMTA*, 2007, pp. 190–195.

[4] C. Esposito, A. Gifuni, and S. Perna, "Measurement of the antenna phase center position in anechoic chamber," *IEEE Antennas and Wireless Propagation Letters*, vol. 17, no. 12, pp. 2183–2187, 2018.

[5] N. Basta and M. Erić, "Method for estimation of electrical distance between antennas based on music-type algorithm," in *7th International Conference on Electrical, Electronics and Computing Engineering (IcETRAN 2020)*, 2020, pp. 1–4.

[6] P. Bidigare, S. Pruessing, D. Raeman, D. Scherber, U. Madhow, and R. Mudumbai, "Initial over-the-air performance assessment of ranging and clock synchronization using radio frequency signal exchange," in *2012 IEEE Statistical Signal Processing Workshop (SSP)*, 2012, pp. 1–4.

[7] G. Tang, X. Qu, F. Zhang, X. Zhao, and B. Peng, "Absolute distance measurement based on spectral interferometry using femtosecond optical frequency comb," *Optics and Lasers in Engineering*, vol. 120, pp. 71–78, 2019.

[8] N. Vukmirović, M. Janjić, P. M. Djurić, and M. Erić, "Position estimation with a millimeter-wave massive MIMO system based on distributed steerable phased antenna arrays," *EURASIP Journal on Advances in Signal Processing*, vol. 2018, no. 1, p. 33, 2018.

[9] N. Vukmirović, M. Erić, M. Janjić, and P. M. Djurić, "Direct wideband coherent localization by distributed antenna arrays," *Sensors*, vol. 19, no. 20, p. 4582, 2019.

[10] R. Di Taranto, S. Muppirisetty, R. Raulefs, D. Slock, T. Svensson, and H. Wymeersch, "Location-aware communications for 5G networks: How location information can improve scalability, latency, and robustness of 5G," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 102–112, 2014.

[11] A. J. Weiss and E. Weinstein, "Fundamental limitations in passive time delay estimation – part I: Narrow-band systems," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, no. 2, pp. 472–486, 1983.

[12] E. Weinstein and A. J. Weiss, "Fundamental limitations in passive time delay estimation – part II: Wide-band systems," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 32, no. 5, pp. 1064–1078, 1984.

Resolvability of Transmitters in Coherent Direct Localization

Nenad Vukmirović, Miloš Janjić, and Miljko Erić

***Abstract*—Coherent direct localization promises high accuracies, that are especially useful for improving wireless link performance (the location-aided communication concept). The focus of the paper is to analyze the spatial resolution performance of three different localization algorithms in this category in the context of spectrum sensing – i.e., their ability to successfully resolve multiple transmitters at different positions working in the same band and time interval. Namely, when two transmitters are close to each other, they interfere with the localization process, which can perceive them as a single source (and, therefore, fail to resolve them). We quantify the impact of this interference on the probability of resolution and the localization error for both cooperative and non-cooperative transmitters. The results of simulations show that, even when the distance between the transmitters is lower than the carrier wavelength, given that the inherent ambiguity problem allows, they can be resolved, with a localization error of a small fraction of the wavelength. The resolution rate is extremely high for the algorithm with a priori known waveform (for cooperative transmitters).**

***Index Terms*—Coherent direct position estimation; distributed antenna array; resolvability of multiple transmitters; spectrum sensing**

I. Introduction

THE focus of the paper is an analysis of spatial resolution performance of coherent localization methods in the context of spectrum sensing. Spatial resolution refers to the ability of an algorithm to correctly perceive two signal sources that are close to each other as two different sources and to estimate their positions, based on the received signals.

Resolution of two known waveforms in noise is analyzed in [1], such as two complex sinusoids of similar frequencies. Resolution performance of direction-of-arrival estimation for sources by a collocated sensor array is assessed in [2]. The paper [3] generalizes the resolution analysis to multiple parameters per signal (such as the spatial coordinates of its source) and multiple signals. The impact of blurring on the resolution in image-forming applications is provided in [4].

The authors of [5] discuss different criteria for successful resolution of acoustic sources for different methods. They propose the valley-to-peak ratio (VPR) as a measure of the quality of resolution. Namely, if a localization method has a criterion function whose maxima represent the estimated positions of the sources and there are two sources of equal intensity close to each other, then their maxima and the minimum between them define the VPR. Our paper is based on simulations in which the maxima corresponding to two radio transmitters are searched for starting at their true positions. If the two search instances (one for each transmitter) end at the same point, it is considered that the transmitters have not been resolved in that attempt. We quantify the performance of resolution by the probability of success and the impact on (deterioration of) the position estimation accuracy. If the VPR is low, the noise and interference have a greater chance of making the resolution process fail (we implicitly rely on the VPR for quantification). Additionally, we generalize the analysis to transmitters of different power levels.

II. Problem Formulation

Let us consider a distributed array of M receive (Rx) antennas at known positions $\vec{r}_m$, $m \in \{1, 2, \ldots, M\}$. The Rx antennas are placed in an area where localization of transmitters (Tx) is performed. We analyze the performance of *coherent* localization. This type of localization requires a propagation medium in which the spatial coherence

N. V., M. J., and M. E. are with the University of Belgrade, School of Electrical Engineering, Serbia

N. V. and M. J. are also with the Innovation Center of School of Electrical Engineering, University of Belgrade, Serbia

M. E. is also with Vlatacom Institute, Belgrade, Serbia

This paper is a result of the research supported by the Serbian Ministry of Education, Science and Technological Development.

condition holds for (at least) the line-of-sight (LoS) components of the signals which the Txs transmit and the Rx array receives. This condition, [6], [7], allows us to use the additional information embedded in the carrier phases of the signals, to increase the accuracy, unlike non-coherent methods.

The raw received signals are processed by the system performing the localization (therefore the localization is *direct*). Obviously, the receiving channels (the Rx antennas, the Rx front-ends, and the signal cables between them) need to be time, frequency, and phase synchronized (t-, f-, and φ-sync). This can be achieved by means of hardware calibration or by processing beacon signals from a dedicated anchor (say, a base station). The baseband complex form of the signal each Rx channel m receives is

$$\begin{aligned} u_m(t) =& \eta_m(t) + \sum_{q=1}^{Q} s_m^{(q)}(t) \\ s_m^{(q)}(t) =& A_m^{(q)} e^{-\mathrm{j}\omega_\mathrm{c}\left(t_0^{(q)}+\tau_m^{(q)}\right)} s^{(q)}\left(t - t_0^{(q)} - \tau_m^{(q)}\right), \end{aligned} \tag{1}$$

where $\eta_m(t) \sim \mathcal{CN}\left(0, \sigma^2\right)$ is an independent white Gaussian noise; Q is the number of Txs; $s^{(q)}(t)$ is the waveform of Tx_q; $t_0^{(q)}$ models the lack of t-sync between Tx_q and the Rx system; $\tau_m^{(q)}$ is the propagation time from Tx_q to Rx_m and $A_m^{(q)}$ is the amplitude factor; the carrier phase is in the exponent and $\omega_\mathrm{c} = 2\pi f_\mathrm{c}$ is the carrier frequency. In this model, the frequencies are normalized by the sampling frequency, $\widetilde{f}_\mathrm{s}$, and time values by $\widetilde{f}_\mathrm{s}^{-1}$, e.g. $f_\mathrm{c} = \widetilde{f}_\mathrm{c}/\widetilde{f}_\mathrm{s}$, $t = \widetilde{t}\widetilde{f}_\mathrm{s}$, $\tau_m^{(q)} = \widetilde{\tau}_m^{(q)}\widetilde{f}_\mathrm{s}$, and so on, where the symbol $\widetilde{\ }$ denotes values in physical units (Hz and s). To keep the analysis tractable, we restrict it to the LoS-only scenario. Then the coherence implies that $A_m^{(q)}$ is real valued.

Specifically, we are interested in analyzing the ability of localization algorithms to distinguish between different Txs which transmit in the same band and time interval (the ability to resolve them successfully) even if they are close to each other. A localization algorithm produces an estimate, $\widehat{\vec{r}^{(q)}}$, of the true (and unknown) location $\vec{r}^{(q)}$ of Tx_q. If, say, $\widehat{\vec{r}^{(1)}} = \widehat{\vec{r}^{(2)}}$ (to within the numerical precision), than the algorithm has failed to resolve the different transmitters Tx_1 and Tx_2. We wish to find the minimum distance between them at which they are still resolvable.

III. The Method

Unlike a system with a single classical (collocated) antenna array, which can estimate the direction of arrival of an incoming radio signal, a system with antennas distributed around the Tx area can estimate their positions, even if they are not t-synchronized with it. We perform Monte-Carlo simulations of such a scenario with $Q = 2$ Txs. We cover the inside of the Rx array aperture by a discrete set of nominal Tx points (the Tx grid). This allows us to average the results over the space. In each simulation run, the positions of Tx_1 and Tx_2 are generated with a specified distance between them and a random orientation at one of the Tx grid points.

A direct localization algorithm in this paper has a criterion function g defined over the area of interest. The only difference between g and a cost function is that a cost function is searched for its minima, whereas g is searched for its maxima. The search is initialized for each of the two Txs at its true location, $\vec{r}^{(q)}$, and it follows the gradient of g to find the maximum, which is the estimate of that Tx's location, $\widehat{\vec{r}^{(q)}}$. Multiple runs are performed at each Tx grid point to achieve a desired statistical sample size. Successful Tx resolutions are counted and the squared Euclidean distance between the estimated and true location of a Tx, $\|\widehat{\vec{r}^{(q)}} - \vec{r}^{(q)}\|^2$, is averaged. Thus, we obtain an estimate of the probability of resolution and the root-mean-square-error (RMSE) of localization for that algorithm.

We use the ML-KS (maximum likelihood – known sequence), ML-US (ML – unknown sequence), and SCM-MUSIC (steered-covariance-matrix multiple-signal classification) algorithms from [6] as representatives of coherent algorithms. ML-KS has stricter requirements for the Tx than the other two. It requires that the modulator in the Tx is coupled with its D/A converter so that the carrier phase is 0 at $t = 0$ (on the local time axis) for each processed signal segment. ML-KS also needs to know the Tx's waveform. This is suitable for localization of cooperative Txs, such as user terminals (UT) in a wireless network, where the base stations allocate training waveforms for the UTs and also perform localization.

ML-US and SCM-MUSIC impose no such restrictions and are suitable even for non-cooperative radio sources. For them, deviations in the modulator phase coupling and carrier frequency can be considered as a part of the Tx waveform itself (since it is unknown to the Rx system, anyway).

A. Grating Lobes

Coherent algorithms suffer from the (integer wavelength) ambiguity problem. It can be intuitively explained like this. If the system performs distributed beamforming in the downlink, than there might appear spots in the area other than the UT antenna location where the electric field vector also has an increased intensity. Localization based on uplink signals would then have high lobes in its criterion function at those spots (the sidelobes), not only at the true UT location (the main lobe).

When analyzing Tx resolvability, we have to consider not only the distance between the Tx_1's main lobe and that of the Tx_2, but also to the closest high sidelobe (the closest grating lobe) of Tx_2. This increases the chance the Txs will interfere with each other. However, if they are moving, it is expected that the overlapping of the lobes will happen only for very short periods of time, so that they would be resolvable most of the time. This definitely seems like an important topic for future research.

IV. Simulation Results

Let us define SNR_0 as the signal-to-noise ratio (SNR) of a Tx's signal in a channel whose Rx antenna would be 1 m away from the Tx. We performed Monte-Carlo simulations with $Q = 2$ Txs, where we kept the SNR_0 of Tx_2 at 30 dB. The (power) level of Tx_1 was 0 dB, -5 dB, and -15 dB relative to Tx_2. The Rx array had 5 antennas at (x, y) coordinates $(-2.195, -1.243)$, $(0.177, -2.641)$, $(2.961, -1.056)$, $(2.534, 2.206)$, and $(-2.18, 2.237)$ in [m]. Each (Tx and Rx) antenna was assumed to have an omnidirectional radiation pattern in the plane of the array. The area inside the array's aperture was covered by a Tx grid with 28×28 points. For each point we performed $K = 3$ simulation runs. In each run, Tx_1 was placed at the corresponding Tx grid point and Tx_2 was placed randomly (with uniform distribution) on a circle centered at Tx_1 with the radius equal to the given distance between Tx_1 and Tx_2, denoted by $d_{12} = \|\vec{r}^{(1)} - \vec{r}^{(2)}\|$. The waveform of each Tx in each run was generated based on a new independent realization of a random complex Gaussian sequence of $N = 256$ samples, with $\widetilde{f}_c = 1$ GHz and $\widetilde{f}_s = 10$ MSample/s.

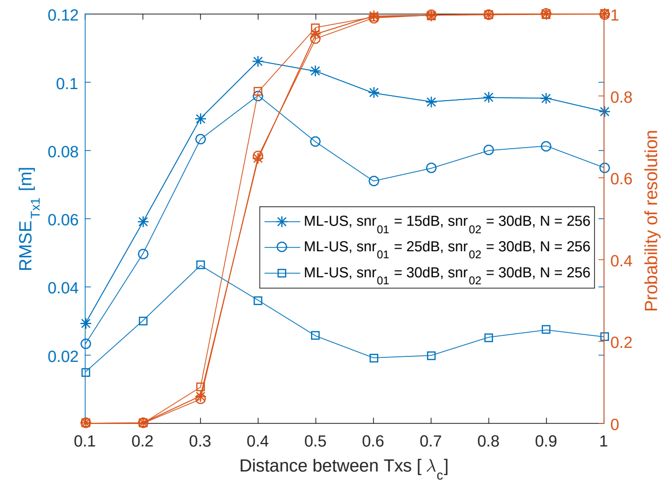


Fig. 1. Probability of resolution and RMSE vs. d_{12} for the ML-US algorithm.

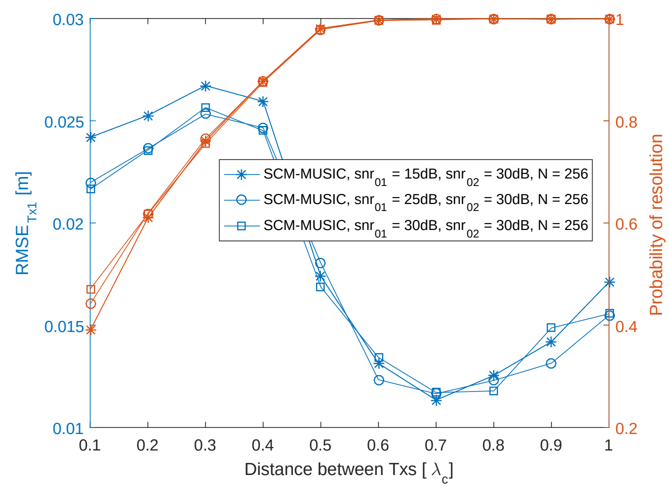


Fig. 2. Probability of resolution and RMSE vs. d_{12} for the SCM-MUSIC algorithm.

The results of simulations for the ML-US algorithm vs. the distance between the Txs (given in carrier wavelengths, λ_c) are shown in Fig. 1. The algorithm resolves the Txs in most cases when $d_{12} \geq 0.5\lambda_c$ and fails in most cases when $d_{12} \leq 0.3\lambda_c$. The RMSE for Tx_1 (the Tx with lower or equal power) is approximately in the range 2 cm – 10 cm when the Txs are successfully resolved. Then the RMSE deteriorates with decreasing d_{12}.

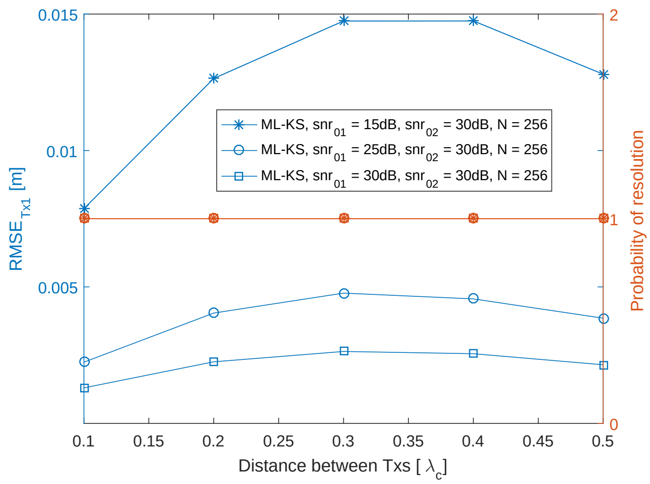


Fig. 3. Probability of resolution and RMSE vs. d_{12} for the ML-KS algorithm.

Note that the RMSE appears to improve for very low d_{12}, but that is just a consequence of Tx_1 location estimate being close to Tx_2, which itself is closer and closer to Tx_1 (there is no actual improvement).

The results for SCM-MUSIC are depicted in Fig. 2. This algorithm has better resolvability in the low-d_{12} region (below $0.5\lambda_c$) than ML-US. Furthermore, all three RMSE curves are similar to the ML-US curve for the best-case scenario of the given three regarding the difference in levels between Tx_1 and Tx_2 (when they transmit at an equal power), so we conclude that SCM-MUSIC is robust with respect to this difference. Better performance of this algorithm is expected, since it is based on a high-resolution algorithm (the generic MUSIC), although, this comes at the cost of higher numerical complexity.

The results for ML-KS are shown in Fig. 3. We can see that it has an extremely high resolvability across the d_{12} range (nearly 1), owing to the fact that the sequences (and waveforms, as well) of the Txs are nearly orthogonal with high probability (they are independent random vectors). We also see that the RMSE is quite low – in most cases it is below 1 cm and it is not affected much by d_{12}. This comes at a price of increased numerical complexity and the reduced scope of applications (as explained in the previous text) – the algorithm has to know the Tx's sequence, so it is mostly for cooperative applications.

One direction for future research is analyzing the impact of different levels of orthogonality between the Txs' waveforms on localization performance. Another is optimization of the Rx antenna array's geometry to suppress the ambiguity problem, effectively reducing the chance the Txs would interfere with each other in the localization process. It would also be interesting to quantify the effect of the ambiguity on localization and tracking, when the Tx is moving.

V. Conclusion

In this paper we presented an analysis of coherent localization performance of multiple transmitters in the same band and time interval, for three different algorithms. The SCM-MUSIC algorithm performs better than ML-US in unfavorable conditions and has the same scope of applications, but at a higher numerical cost. The ML-KS has the best performance, but at a higher numerical cost and it is usually restricted to localization of cooperative transmitters. All in all, each of the analyzed algorithms have a localization error that is a small fraction of the carrier wavelength (as long as the ambiguity problem does not cause them to fail), despite the fact that the transmitters interfere with each other.

References

[1] A. Amar and A. J. Weiss, “Fundamental limitations on the resolution of deterministic signals,” *IEEE Transactions on Signal Processing*, vol. 56, no. 11, pp. 5309–5318, 2008.

[2] J.-P. Delmas, *Performance bounds and statistical analysis of DOA estimation*. Academic Press library in signal processing, Elsevier, 2015.

[3] M. N. El Korso, R. Boyer, A. Renaux, and S. Marcos, “Statistical resolution limit for multiple parameters of interest and for multiple signals,” in *Proc. of International Conference on Acoustics, Speech, and Signal Processing, (ICASSP10)*, 2010, pp. 3602–3605.

[4] G. M. Gasimova, R. C. Masmaliyeva, and G. N. Murshudov, “Resolution and resolvability in one, two and three dimensions,” *Journal of Life Sciences and Biomedicine*, vol. 1, no. 1, pp. 29–36, 2019.

[5] D. Fernandez Comesaña, K. R. Holland, and E. Fernandez-Grande, “Spatial resolution limits for the localization of noise sources using direct sound mapping,” *Journal of Sound and Vibration*, vol. 375, pp. 53–62, 2016.

[6] N. Vukmirović, M. Erić, M. Janjić, and P. M. Djurić, “Direct wideband coherent localization by distributed antenna arrays,” *Sensors*, vol. 19, no. 20, p. 4582, 2019.

[7] N. Vukmirović, M. Erić, and P. M. Djurić, “Performance limits of direct wideband coherent 3d localization in distributed massive mimo systems,” *Sensors*, vol. 21, no. 10, p. 3401, 2021.

Performance analysis of LDPC and Polar codes for message transmissions over different channel models

Darija Čarapić, Mirjana Maksimović and Miodrag Forcan

***Abstract*— The realization of a wireless communication system that will meet the demands of the modern world in terms of fast, secure, reliable, and cost-effective information exchange, is a challenging task. Having in mind that the transmission of data takes place in an imperfect channel environment where noise, fading, and interference are present, the achievement of timely communication with a minimum of errors during data transfer requires the right choice of the channel coding scheme. Channel coding is a fundamental component of the communication system and is intended to ensure that the information received is the same as the sent one. Two coding schemes are available in the fifth generation of mobile communications (5G): Low-Density Parity-Check (LDPC) codes for coding user information and Polar codes for coding control information. This paper presents a comparative simulation study of LDPC and Polar codes for message transmissions over different channel models (Additive White Gaussian Noise (AWGN), Rician, and Rayleigh). The Bit Error Rate (BER) performance of these codes was reviewed for all three channel models. The simulations considered variable message sizes and code rates for LDPC and Polar codes, different modulation patterns for LDPC codes, and different decoding schemes for Polar codes. The results of the simulations showed the performances of the LDPC and Polar codes in the case of channel models: AWGN with no fading and AWGN with fading. The LDPC codes have been superior in the case of long messages and the Polar codes have been more efficient in the case of short messages, hence justifying the use of both LDPC and Polar codes within the 5G.**

***Index Terms*— LDPC, Polar, BER, AWGN, Rician, Rayleigh**

I. Introduction

A communication medium is prone to errors due to random noise, interference, fading, device impairments, etc. Channel impairments lead to the corruption of the original data flow, so the data on the receiving side is not the same as the data that was sent. To correct the errors made during data transmission, channel coding is applied. This means that on the transmitter side, the original data flow is subjected to a series of algorithmic operations (channel encoding). On the receiver side, channel decoding is done by applying other operations set to correct errors. It is obvious that the choice of an adequate coding scheme is of paramount importance for the rapid and reliable transmission of data. Enhanced flexibility, low computation complexity, low latency, low cost, and high reliability are desired features for the coding scheme [1, 2].

In contrast to the previous generations of mobile communication systems (3G and 4G), which used Turbo code as a channel coding technique, the fifth generation of mobile communications (5G) introduces two coding techniques: Low-Density Parity-Check (LDPC) and Polar codes. The reasons Turbo code is not chosen for the 5G are numerous iterations and a significant delay in decoding. As such, Turbo codes are unable to meet the demands of 5G networks in terms of high speed and low delay. LDPC and Polar codes are chosen as 5G channel coding standards due to their features and because the unique coding technology can no meet the needs of all scenarios and users in 5G. LDPC codes have better band utilization, and they perform better for longer block lengths in comparison with Polar codes that are superior for shorter block lengths. Furthermore, compared to other codes, LDPC codes have better decoding latency, throughput, and implementation while Polar codes are simple to implement, can reach the channel capacity and have low encoding and decoding complexity [3]. According to 3rd Generation Partnership Project (3GPP) 5G New Radio (NR) standardization, LDPC coding scheme is proposed for use in data channels while Polar codes are applied in control channels [4].

This paper presents an attempt to simulate the performances of LDPC and Polar codes in the case of different channel models (Additive White Gaussian Noise (AWGN), Rician fading + AWGN, and Rayleigh fading + AWGN) in order to summarize the advantages and disadvantages of 5G channel coding techniques. A comparative simulation study is performed for variable message sizes and code rates. Moreover, different modulation schemes were analyzed in the case of LDPC codes as well as different decoding schemes for Polar codes.

Therefore, the paper is organized in the following manner. A short overview of communication channel models is presented in Section II. Section III presents 5G channel coding schemes while the results of simulations (Bit Error Rate (BER) vs Signal-to-Noise Ratio (SNR) graphs) are given in Section IV. A summary of the performed research and directions for future research are provided in the Conclusion.

II. Communication channel models

The theory of communication is often based on the assumption that the transmitted signals are distorted by a certain noise. The most commonly used noise assumption is additive, white, with Gaussian-distributed values. The

Darija Čarapić is with the m:tel a.d. Banja Luka, Bosnia and Herzegovina, (e-mail: darija_carapic@yahoo.com).

Mirjana Maksimović is with the Faculty of Electrical Engineering, University of East Sarajevo, Vuka Karadžića 30, 71123 East Sarajevo, Bosnia and Herzegovina (e-mail: mirjana.maksimovic@etf.ues.rs.ba).

Miodrag Forcan is with the Faculty of Electrical Engineering, University of East Sarajevo, Vuka Karadžića 30, 71123 East Sarajevo, Bosnia and Herzegovina (e-mail: miodrag.forcan@etf.ues.rs.ba).

model of AWGN channel is used to simulate the influence of naturally occurring random signals on wireless signals. It typically represents the background noise, amplifier noise in the transceivers, and signals from other communication systems in the frequency bands [5].

It is important to highlight that fading, frequency selectivity, interference, non-linearity, or dispersion are not considered in the AWGN channel model what makes it unrepresentative for most wireless connections. A more realistic scenario of wireless channels is the availability of multiple paths between transmitters and receivers. These routes can be direct or formed through reflection, diffraction, or scattering. In this case, the receiver receives distinct copies of the sent signal (with variable attenuation, delay, or phase shift) [6, 7]. Rayleigh distribution (for the scenario when no Line of Sight (LoS) component is present) and Rician distribution (when LoS component is present) are often used to represent multipath fading in wireless communication systems. In the Rician channel model, a receiver receives a direct LoS signal from the source in addition to the other non-LOS components. Rayleigh fading is considered as a special case of Rician. The receiver cannot receive any LoS signal directly from the source. All incoming signals are diffracted, reflected, or diffused. Rayleigh fading channel model is the right choice when there are several objects in the surroundings that scatter the signal before it reaches the receiver [8].

III. 5G Channel coding schemes

Data transmission over the wireless communication link that is subject to interference and fading can result in the data received is different than the data sent. In order to overcome this problem, additional information is added to the data sent by the transmitter. At the receiver side, complex schemes that need sophisticated algorithms decode this information and retrieve the original data. This process is called Forward Error Correction (FEC) or channel coding and has an immense role in increasing the performance of wireless communication systems.

3GPP 5G NR brings requirements to channel coding at a completely new level. 5G is designed to serve a wide range of applications: enhanced mobile broadband (eMBB), ultra-reliable and low-latency communication (URLLC), and massive machine-type communication (mMTC) [9].

5G introduces the two capacity-achieving channel coding schemes, LDPC and Polar codes, not used in earlier generations of mobile communication systems. Hence, in the 5G communication systems, the channel coding has been separated into channel coding of user information (LDPC coding) and channel coding of control information (Polar coding).

A. LDPC codes

LDPC codes are a sort of linear block code first proposed by Gallager in the early 1960s [10] and rediscovered in the late 1990s thanks to MacKay [11, 12]. The name of these codes comes from the fact that their parity-check matrix is largely zeros (0s) with a minor number of ones (1s). LDPC codes can be described via matrices or represented graphically (Tanner graphs). There are two basic graphs and eight sets of lifting sizes in the 5G NR, hence allowing a variety of block lengths and coding rates. Graph structure sparsity has a strong influence on the algorithmic efficiency of LDPC codes. According to [13], it is not difficult to construct effective LDPC codes. Actually, the random codes have a high probability of success. The issue is that the encoding complexity of these codes is typically quite high. Decoding algorithms for LDPC codes are relatively simple and practical. The belief propagation algorithm, the message passing algorithm, and the sum-product algorithm are the most commonly used [13, 14]. The main advantages of LDPC codes are good block error performance, error floors in much lower BER values, the ability to achieve good error performance without requiring interleavers, and an iterative-based decoding process [15]. Because LDPC codes perform close to a channel's Shannon limit only for long block lengths, they are well suited for use in 5G NR for user data transmission.

3GPP specifies LDPC coding chains for the 5G NR downlink shared transport channels (DL-SCH) and uplink shared transport channels (UL-SCH) [4]. In 5G NR, data are transmitted in units referred to as transport blocks. The term "transport blocks" refers to data sent from the Medium Access Control (MAC) layer to the physical layer. A transport block goes through the processing steps shown in Fig. 1 before transmitting to the Physical Downlink Shared Channel (PDSCH) for scrambling, modulation, layer mapping, and resource/antenna mapping.

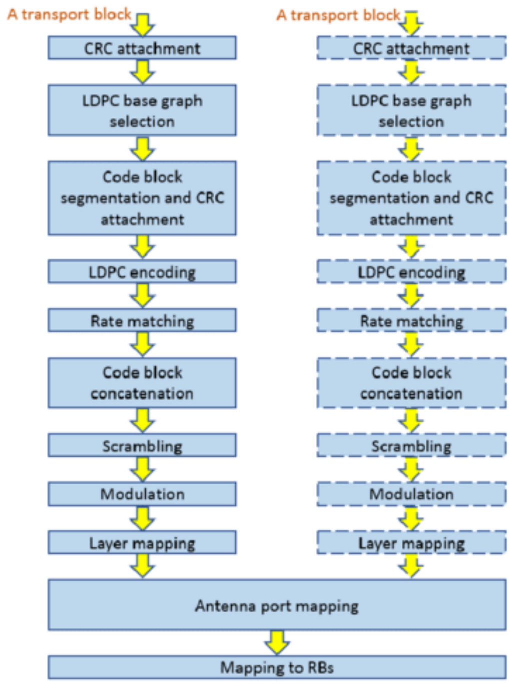


Fig. 1. 5G NR PDSCH physical layer processing [16]

PDSCH is used for a variety of data transmissions, including downlink user data, user equipment specific higher layer information, system information blocks, and paging [16]. Depending on the link conditions, the PDSCH employs an adaptive modulation format (QPSK, 16-QAM, 64-QAM, and 256-QAM). It also employs a flexible coding scheme. When these elements are combined, it results in a flexible coding and data rate [17]. The Physical Uplink Shared Channel (PUSCH) is the PDSCH's counterpart. PUSCH transmits an uplink shared channel (UL-SCH) and its higher mapped channel data. The physical layer of an uplinked transport block is processed similarly to that of a downlinked transport block (Fig. 2). PUSCH also has a very adaptable format. Frequency resources are allocated using blocks of resources as well as a flexible modulation (pi/2-BPSK, QPSK, 16-QAM, 64-QAM, and 256-QAM) and coding scheme depending on the link conditions [16, 17].

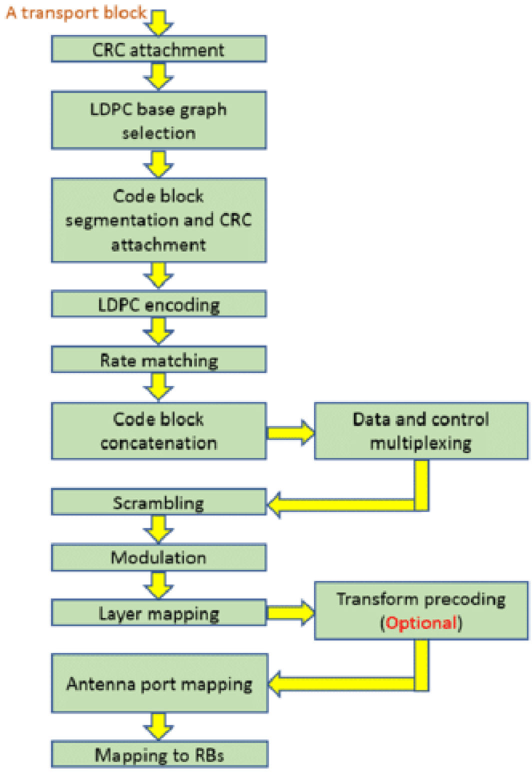


Fig. 1. 5G NR PUSCH physical layer processing [16]

B. Polar codes

Polar codes are another official channel coding scheme accepted in 5G standardization. The idea behind Polar codes, introduced by Arikan [18] in 2009, is to subdivide the original channel into a number of virtual channels, each of which is purely noisy or noiseless. By sending data over noise-free channels while the fixed bits, which are known at both the encoder and decoder, are sent over noisy channels, nearly error-free transmission is possible [19]. Polar codes are applied as channel codes for 5G NR control channels (Physical Downlink Control Channel (PDCCH) and Physical Uplink Control Channel (PUCCH)) where blocks of information are small, and the Hybrid Automatic Repeat Request (HARQ) is not used [14].

As indicated by the name, the 5G PDCCH transports downlink control information (DCI). Its main function is to schedule downlinking transmissions on the PDSCH as well as uplinking data transmissions on the PUSCH. With the exception of small data packets, the PDCCH employs QPSK as a modulation format and Polar coding as a coding scheme [17]. 5G NR PDCCH physical layer processing steps are shown in Fig. 3. It is important to point out that for downlink transmission, the size of the transport block is limited to 36-164 bits (interleaver limit).

The main purpose of PUCCH is to carry Uplink Control Information (UCI) such as HARQ feedback, channel state information, and scheduling request. The PUCCH employs BPSK or QPSK as a modulation format. As a coding scheme, PUCCH uses Polar coding. In the uplink transmission, the size of the transport block is limited to 31-1024 bits. If the payload size is greater than or equal to 1013 bits, code block segmentation for uplink is carried out [16].

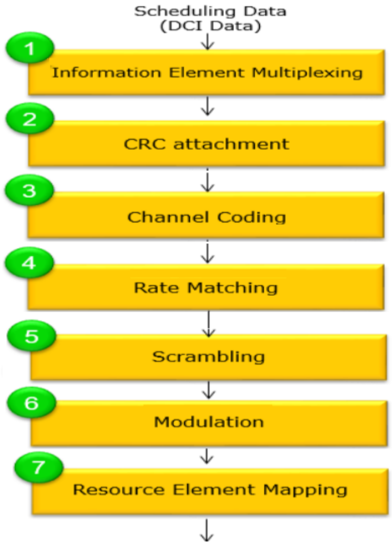


Fig. 3. 5G NR PDCCH physical layer processing [20]

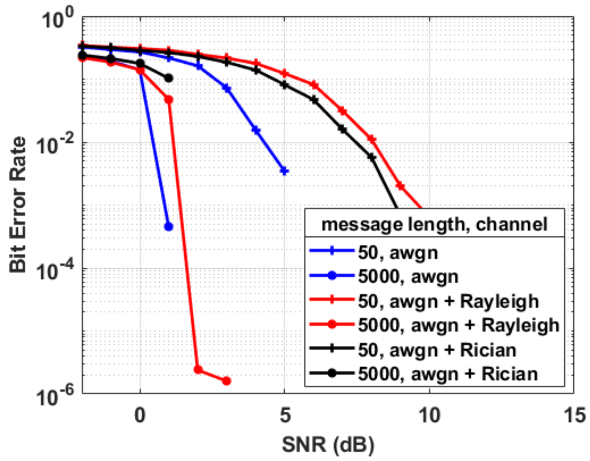
As a starting point for Polar decoding, Arikan proposed decoding using successive cancellation (SC). Despite the advantage of low complexity, it is not adequate for block lengths ranging from short to medium. This problem can be solved by using the successive cancellation list (SCL) decoding algorithm. The SCL decoder uses a list size parameter L (the number of decoding paths that are most likely to be retained) to decode the input bits one by one. In 5G error-correction performance evaluations, Cyclic Redundancy Check (CRC)-aided SCL has been used. Although SCL's effectiveness increases as the list size parameter L increases, so does its implementation complexity. SCL becomes SC when L is set to 1 [14, 21].

IV. 5G CHANNEL CODING SCHEMES' BER PERFORMANCE

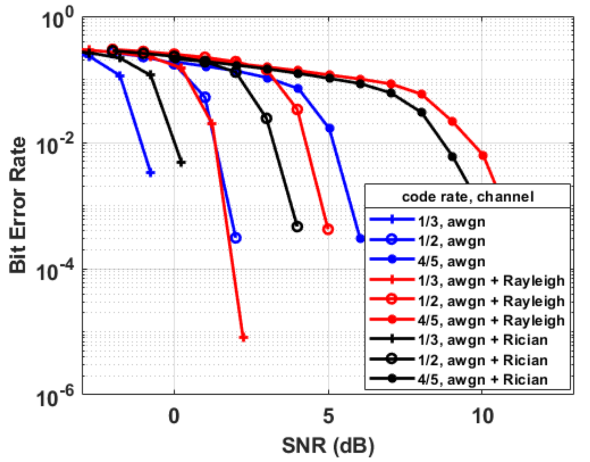
The MATLAB R2020a software package [22] was used to simulate the 5G physical communication layer with LDPC and Polar coding schemes. The downlink and uplink coding schemes are implemented in accordance with 3GPP regulations [4]. The performance of BER is analyzed for different message lengths (50, 500, 5000, and 50000 bits) and variable code rates using different communication channels: AWGN, Rician, and Rayleigh. The variances of AWGN channel model are estimated using SNR values. In the case of fading presence, a fading channel block was accompanied by the AWGN channel block that had previously been used. Fading channel property values used in simulations are: sampling rate: 10^5 [s]; path delays: 0, 10^{-7} and 10^{-5} [s] (for the outdoor environment); average path gains: 0, -3, and -3 [dB]; a maximum Doppler shift of 0; and the Rician K factor is set to 3 (K is the power of the LoS component divided by the power of the scattered components) [14, 23]. BER is calculated and plotted against the SNR. Each simulation was performed for 500 frames and continued until the BER of 10^{-6} is achieved. QPSK has been used in all simulations [23].

A. LDPC codes

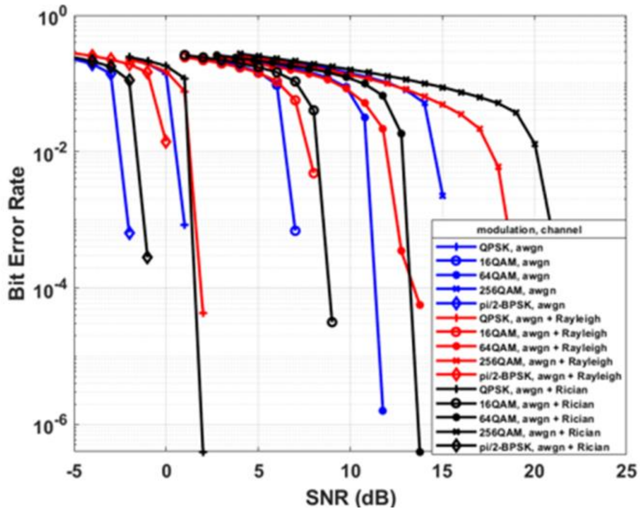
Fig. 4 shows BER vs SNR graph for LDPC coded uplink transmission (quite similar simulation results are obtained for downlink transmission). Fig. 4. a) presents BER performance for messages of different lengths (50 bits and 5000 bits). The selected code rate is ½. As can be seen, for the longer message LDPC shows better performance. BER performance for different code rates is given in Fig. 4. b). The simulation is performed for 500 bits long message in the uplink direction. Simulation results demonstrate that the lowest code rate means the longest coded word and better BER performance. Since LDPC codes show better performance for longer words, Fig. 4. c) presents BER vs SNR graph for 5000 bits long message in case of different modulation techniques. In uplink directions, pi/2-BPSK, QPSK, 16-QAM, 64-QAM, and 256-QAM modulation schemes are supported while the opposite direction supports QPSK, 16-QAM, 64-QAM, and 256-QAM modulation schemes. Obtained results show the best performance for pi/2-BPSK and QPSK modulation schemes. Since pi/2-BPSK is not supported in both directions, all other simulations have been performed using QPSK modulation scheme.



a) variable message length (code rate=1/2);



b) variable code rates (message length =500 bits);



c) different modulation techniques (message length =5000 bits)

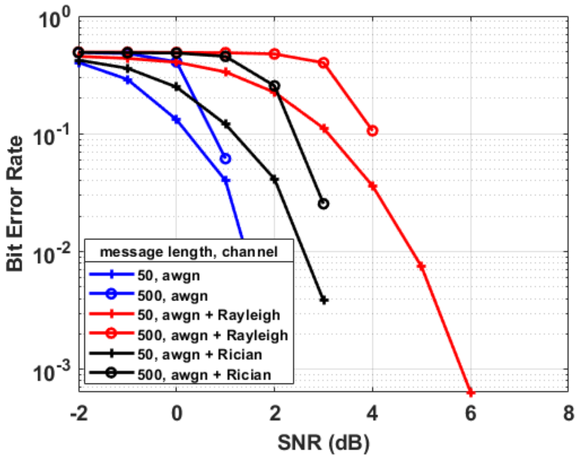
Fig. 4. BER performance for LDPC coded uplink data transmission

The common feature of all three graphs in Fig. 4 is that the AWGN channel model achieves the best performance, and the Rayleigh channel model achieves the worst. This is because the performance of the BER is considerably improved in the case of low SNR, but not in the case of high SNR. White Gaussian noise dominates the BER error when SNR is low. In this case BER performance can be improved by increasing SNR. However, when the SNR is high, the phase estimation error dominates the BER error. Simply increasing the SNR will not improve BER performance in this case [8].

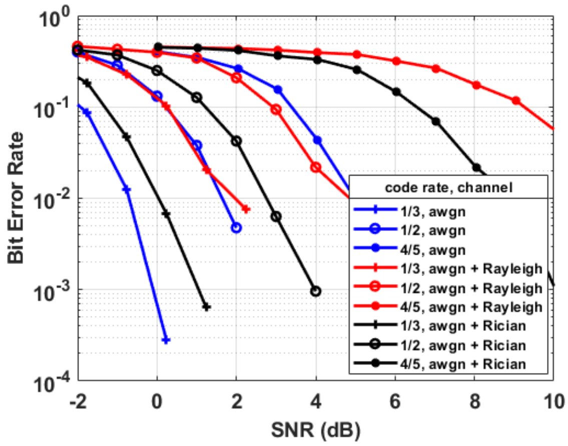
B. Polar codes

A similar simulation study was also conducted in the case of Polar codes. Fig. 5 shows the BER vs SNR graphs for uplink transmission of Polar coded messages. Fig 5. a) considers different message lengths (50 bits and 500 bits). Uplink transmission of a longer message is not supported according to 3GPP (the maximum size of input length is 1023 bits). Among selected message lengths, BER performance simulation in downlink directions can be done only for a 50 bits long information message since the maximum input length is 164 bits. Fig. 5. a) results have been achieved for code rate = 1/2. As can be seen, opposite to LDPC, Polar codes show better performance for shorter messages. Since Polar codes show better performances for a shorter message, Fig. 5. b) shows the BER performance for 50 bits long message in case of different code rate values. The best results are achieved for the lowest code rate. Fig. 5 c) shows the simulation results for the CRC-aided SCL decoding algorithm. Variable list sizes *L* (4, 16) have been considered for code rate = 1/2. Results confirm that the larger list size *L* means enhanced Polar coding performance (lower error rate), but implementation complexity increases with higher *L* values.

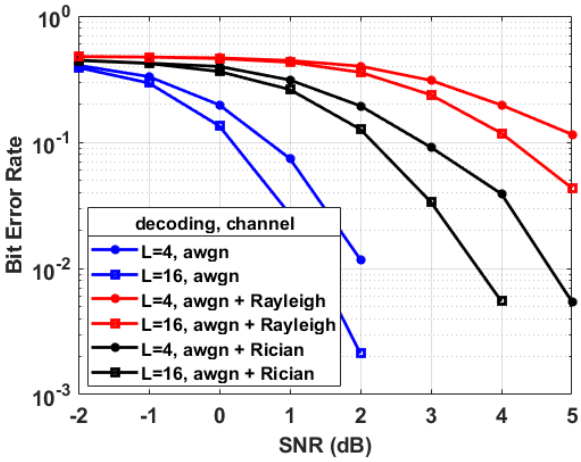
The same as in the LDPC case, the best simulation results are achieved for the AWGN channel. The Rician channel model outperforms the Rayleigh use case in terms of BER.



a) variable message length (code rate=1/2);



b) variable code rates (message length =500 bits);



c) different L values (message length =50 bits)

Fig. 5. BER performance for Polar coded uplink data transmission

C. LDPC vs Polar codes

Fig. 6 demonstrates a comparison of LDPC and Polar coding techniques in both directions. Fig 6. a) shows BER vs SNR graph for uplink data transmission of 50- and 500 bits messages. Simulation results are achieved for code rate = 1/2, and show that LDPC codes have better performances for longer messages while Polar codes are superior in the case of short messages. Fig. 6 c) considers downlink transmission and knowing that the information message length, in this case, is limited to 164 bits, comparative analysis has been performed only for 50 bits long message. Polar codes outperform LDPC codes because messages are shorter. Fig 6. b), and d) present BER performance for 50 bits long message in case of variable code rates in both directions. For 50 bits long message, Polar codes are better than LDPC in both uplink and downlink directions and at lower code rates show better performance. Following the results from Fig. 6 a), it is evident that the LDPC would have better performance in the uplink transmission of a longer message. In all simulations, the best results are achieved for the AWGN channel model. When comparing fading channel models, the Rician channel model shows better performance in comparison with the Rayleigh channel model.

V. Conclusion

Recognizing the importance of selecting the right channel coding scheme to ensure fast and error-free transmission of data, this paper presents an attempt to conduct a comparative simulation study of 5G channel coding techniques. The performance of the LDPC and Polar codes is studied for three channel models (AWGN, Rician, and Rayleigh) taking into consideration the size of the messages and variable code rates. Performance of 5G channel coding techniques is measured by their ability to correct errors at a given SNR. The results have confirmed LDPC codes' superiority for longer messages and Polar codes' superiority for short messages. In this way, using LDPC codes in data channels and Polar codes in control channels is justified. In addition, for lower SNR values, Gaussian noise dominates the BER error and just improving the SNR, BER error can be improved. For higher SNR values, this cannot be performed as the phase estimation error dominates the BER error. That is why the simulation results have shown that BER performance of the AWGN channel model is better than for fading channel models. Apart from measuring the performance of a channel coding techniques via BER vs SNR graphs, it is also important to analyze the maximum possible throughput, latency and the resources and power consumption. This is the direction of our future research.

References

[1] M. Maksimovic and M. Forcan, "Application of 5G Channel Coding Techniques in Smart Grid: LDPC vs. Polar coding for Command Messaging," *7th International Conference on Electronics, Telecommunications, Computing, Automatics and Nuclear Engineering - IcETRAN* 2020, TEI 1.3.1-1.3.6.

[2] D. Čarapić and M. Maksimović, "A Comparison of 5G Channel Coding Techniques – LDPC and Polar codes", *IJEEC*, Vol.4., No. 2. (2020), pp. 71-82.

[3] S. Cheng, "Comparative Study on 5G Communication Channel Coding Technology", 3rd International Conference on Mechatronics Engineering and Information Technology (ICMEIT 2019), *Advances in Computer Science Research*, Vol. 87, pp. 74-77, 2019.

[4] 3GPP TS 38.212 V15.8.0 (2019-12) 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; NR; Multiplexing and channel coding (Release 15)

[5] S. Host, "Gaussian Channel," in *Information and Communication Theory* , IEEE, 2019, pp.237-263, doi: 10.1002/9781119433828.ch9

[6] H. Zarrinkoub, *Understanding LTE with MATLAB: from mathematical foundation to simulation, performance evaluation and implementation*, John Wiley & Sons, Ltd, 2014

[7] M. K. Samimi, G. R. Maccartney, S. Sun, and T. S. Rappaport, "28 GHz millimeter-wave ultrawideband small-scale fading models in wireless channels," in *Proceedings of the 83rd IEEE Vehicular Technology Conference, VTC Spring 2016*, May 2016.

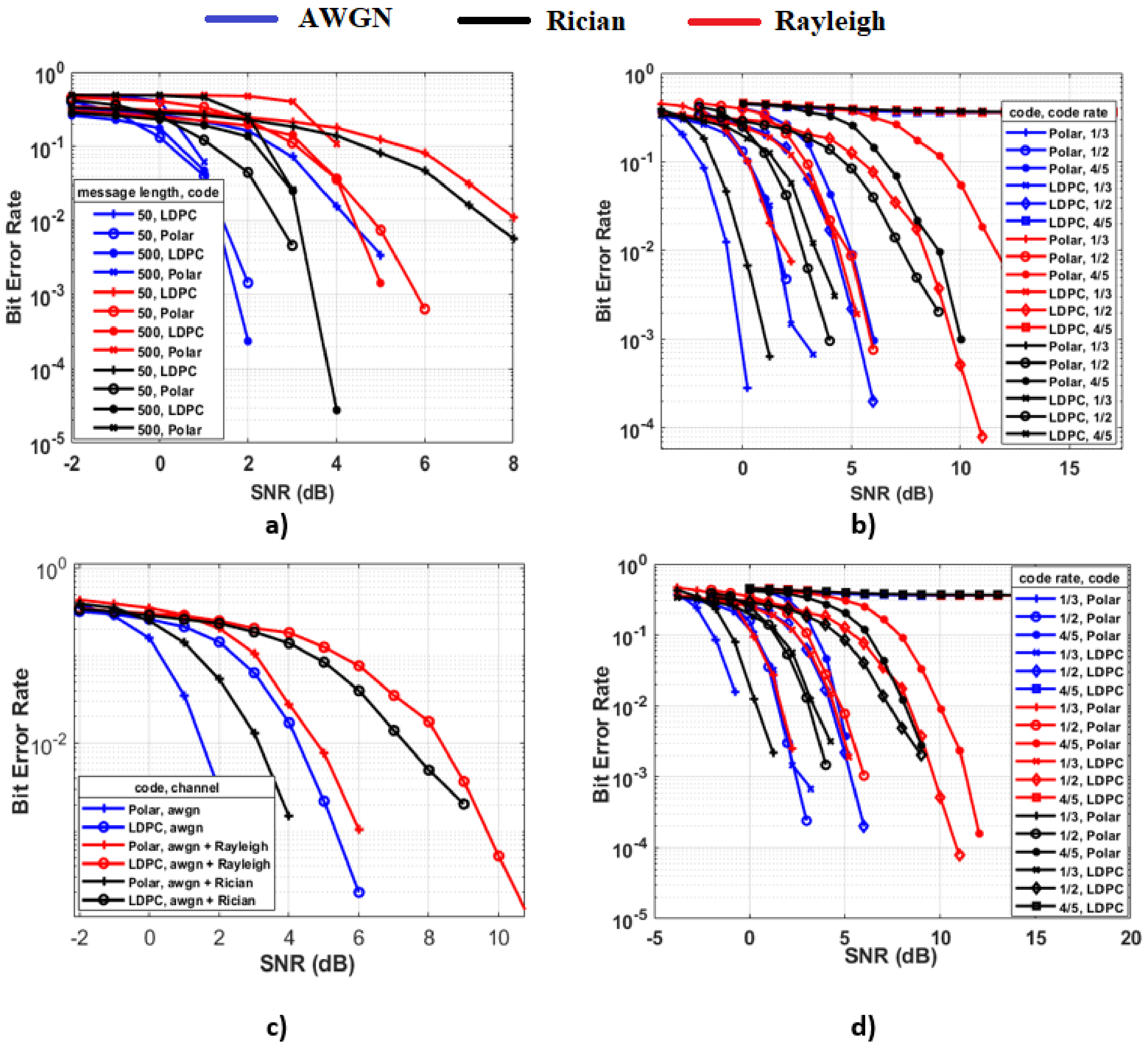


Fig. 6. BER performance – LDPC vs. Polar codes: a) variable message length (code rate=1/2) - uplink; b) variable code rates (message length =50 bits) - uplink; c) message length 50 bits (code rate=1/2) - downlink; d) variable code rates (message length =50 bits) – downlink.

[8] M.S. Chavan, R.H. Chile and S.R. Sawan "Multipath Fading Channel Modeling and Performance Comparison of Wireless Channel Models," *International Journal of Electronics and Communication Engineering*, Vol. 4, Number 2 (2011), pp. 189-203

[9] 3GPP: TR 38.913 Study on Scenarios and Requirements for Next Generation Access Technologies, 2017.

[10] R.G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: MIT Press, 1963

[11] D.J.C. MacKay and R. M. Neal, "Near Shannon limit performance of low density parity check codes." *In Electron. Letters*, Vol.32, August 1996, 1645-1646 (reprinted with printing errors corrected in Vol.33, 457–458)

[12] D.J.C. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Transactions on Information Theory*, 1999, 45(2):399-431

[13] B.M.J. Leiner, *LDPC Codes – a brief Tutorial*, [Online]: Available: http://www.bernh.net/media/download/papers/ldpc.pdf

[14] M. Maksimovic and M. Forcan, "5G New Radio channel coding for messaging in Smart Grid," *Sustainable Energy, Grids and Networks*, Vol. 27, 2021, 100495

[15] I. Develi, and Y. Kabalci, "A comparative simulation study on the performance of LDPC coded communication systems over Weibull fading channels," *J. appl. res. Technol*, vol.14, no.2, 2016, https://doi.org/10.1016/j.jart.2016.04.001

[16] X. Lin *et al*., "G New Radio: Unveiling the Essentials of the Next Generation Wireless Access Technology," in *IEEE Communications Standards Magazine*, vol. 3, no. 3, pp. 30-37, 2019, doi: 10.1109/MCOMSTD.001.1800036.

[17] Electronics Note, *5G Data Channels: Physical; Transport; & Logical*, [Online]: Available at: https://www.electronics-notes.com/articles/connectivity/5g-mobile-wireless-cellular/data-channels-physical-transport-logical.php

[18] E. Arikan, "Channel Polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," in *IEEE Transactions on Information Theory*, 55(7):3051–3073, 2009

[19] K. Niu, K. Chen, J. Lin and Q. T. Zhang, "Polar codes: Primary concepts and practical decoding algorithms," in *IEEE Communications Magazine*, vol. 52, no. 7, pp. 192-203, July 2014, doi: 10.1109/MCOM.2014.6852102.

[20] Sharetechnote, *5G/NR – PDCCH*, [Online]: Available at: http://www.sharetechnote.com/html/5G/5G_PDCCH.html

[21] V. Bioglio, C. Condo and I. Land, "Design of Polar Codes in 5G New Radio," in *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 29-40, 2021, doi: 10.1109/COMST.2020.2967127.

[22] MATLAB, 2020. version R2020a, Natick, Massachusetts: The MathWorks Inc.

[23] D. Čarapić, *Analysis of 5G channel coding techniques*, Master Thesis, University of East Sarajevo, Faculty of Electrical Engineering, June 2021.

ВЕШТАЧКА ИНТЕЛИГЕНЦИЈА
/
ARTIFICIAL INTELLIGENCE

(ВИ/VII)

Rešavanje problema ekonomične raspodele snaga generatora primenom fazorske optimizacije roja čestica

Milena Jevtić, Miroljub Jevtić, Jordan Radosavljević, Sanela Arsić i Dardan Klimenta

***Apstrakt*— Minimizacija troškova goriva i emisije štetnih gasova u termoelektranama podešavanjem izlaznih snaga generatora je jedan od važnih problema u upravljanju elektroenergetskim sistemima. Ovaj problem je poznat kao Combined economic emision dispach (CEED) problem. U ovom radu je za rešavanje CEED problema predložen meta-heuristički algoritam pod nazivom Fazorska optimizacija roja čestica, koji predstavlja unapređenu varijantu Optimizacije roja čestica. Parametri Fazorske optimizacije roja čestica se tokom iteracija automatski podešavaju pa je ovaj algoritam, adaptivni i neparametarski, što je njegova prednost. Performanse predloženog algoritma za rešavanje CEED problema se u radu ocenjuju na standardnom IEEE test sistemu sa 30 čvorova i 6 generatora. Na osnovu dobijenih rezultata utvrđeno je da ovaj algoritam ima bolje karakteristike od algoritama koji su primenjeni u drugim publikovanim radovima za rešavanje CEED problema.**

***Ključne reči* - Combined economic emision dispach; Fazorska optimizacija roja čestica; Upravljanje elektroenergetskim sistemima.**

I. UVOD

Ekonomična raspodela snaga generatora sa istovremenom minimzacijom emisije štetnih gasova (eng. Combined Economic and Emission Dispatch (CEED)) predstavlja podešavanje izlaznih snaga određenog broja generator u termoelektranama, pri zadatom opterećenju i pri zadatim ograničenjima u sistemu, minimizirajući troškove goriva i emisiju štetnih gasova. Funkcije koje opisuju emisiju štetnih gasova i troškove goriva, uzimajući u obzir efekat ventila u elektrani, su nelinearne i nekonveksne tako da je CEED problem u literaturi rešavan metaheurističkim optimizacionim algoritmima koji daju približno rešenje. U publikovanim radovima je predložen veći broj metaheurističkih algoritama u cilju dobijanja što tačnijeg i što bržeg rešenja ovog problema [1], [2], [3]. Brzina i tačnost ovih algoritama utiču na kvalitet softvera, u koje se inkorporiraju, a koji služe za upravljanje emisijom gasova i troškovima goriva u termoelektrani.

U ovom radu se za rešavanje CEED problema predlaže primena jednog od najnovijih algoritama, Fazorske optimizacije roja čestica (eng. Phasor particle swarm optimizacition (PPSO)) [4].

Cilj ovog rada je da se pokaže da PPSO može efikasno da se primeni za rešavanje CEED problema i da daje bolje rezultate u odnosu na druge algoritme koji su u literature primenjivani za rešavanje istog problema.

Milena Jevtić – Tehnički fakultet u Boru, Univerzitet u Beogradu, Vojske Jugoslavije 12, 19210 Bor, Srbija (e-mail: mjevtic@tfbor.bg.ac.rs).

Miroljub Jevtić – Fakultet tehničkih nauka u Kosovskoj Mitrovici, Univerzitet u Prištini, Knjaza Miloša 7, 38220 Kosovska Mitrovica, Srbija (e-mail: miroljub.jevtic@pr.ac.rs).

Jordan Radosavljević - Fakultet tehničkih nauka u Kosovskoj Mitrovici, Univerzitet u Prištini, Knjaza Miloša 7, 38220 Kosovska Mitrovica, Srbija (e-mail: jordan.radosavljevic@pr.ac.rs).

Sanela Arsić – Tehnički fakultet u Boru, Univerzitet u Beogradu, Vojske Jugoslavije 12, 19210 Bor, Srbija (e-mail: saarsic@tfbor.bg.ac.rs).

Dardan Klimenta - Fakultet tehničkih nauka u Kosovskoj Mitrovici, Univerzitet u Prištini, Knjaza Miloša 7, 38220 Kosovska Mitrovica, Srbija (e-mail: dardan.klimenta@pr.ac.rs).

II. CEED MODEL

Funkcija troškova goriva generatora u termoelektrani obično ima kvadratni oblik:

$$F_g\left(P_g\right) = a_g + b_g P_g + c_g P_g^2, \qquad g = 1, 2, \ldots, G \tag{1}$$

gde su: F_g (\$/h) troškovi goriva g- tog generator, P_g (MW) izlazna snaga g-tog generatora, a_g, b_g i c_g koefficienti.

Funkcija $F_g (P_g)$ postaje ne-konveksna kada se uzme u obzir promena snage zbog sekventnog otvaranja ventila u termoelektrani (efekat ventila) [5]:

$$F_g\left(P_g\right) = a_g + b_g P_g + c_g P_g^2 + \left| d_g \sin\left(e_g \left(P_g^{\min} - P_g\right)\right)\right| \tag{2}$$

gde su: d_g i e_g koeficijenti koji se odnose na efekat ventila i P $P_g^{\min}$ donja granična snaga g-tog generatora.

Funkcija koja modeluje emisiju gasova u termoelektrani se predstavlja kao zbir kvadratne i eksponencijalne funkcije izlazne snage generatora [6], [7]:

$$E_g\left(P_g\right) = \alpha_g + \beta_g P_g + \eta_g P_g^2 + \xi_g \exp\left(\lambda_g P_g\right) \tag{3}$$

gde su: E_g (t/h) količina gasova koji se emituju tokom rada g-tog generatora, P_g (MW) izlazna snaga g-tog generatora, i α_g, β_g, η_g, ξ_g *i* λ_g emisioni koeficijenti.

Ako se (1) i (2) kombinuju sa (3), dobija se sledeća funkcija [8]:

$$FE = w\sum_{g\in G} F_g\left(P_g\right) + \left(1-w\right)\gamma\sum_{g\in G} E_g\left(P_g\right) \quad (4)$$

gde su: γ factor skaliranja, w težinski faktor čija vrednost se uzima u granicama $0 < w < 1$, i G je ukupni broj generatora koji se razmatraju, priključenih na system. CEED problem se rešava tako što se izabere faktor w a zatim minimizuje funkcija (4). Izborom gornje granice težinskog faktora, $w = 1$, minimizuje se samo funkcija F_g (P_g), izborom donje granice, $w = 0$, minimizuje se samo funkcija E_g (P_g), dok izbor drugih vrednosti težinskog faktora odgovara istovremenoj minimizaciji troškova goriva i emisije gasova. Faktor skaliranja γ je uveden da bi se funkcija (4), rešavala kao jednociljni optimizacioni problem umesto kao dvociljni.

Minimizacija se vrši za zadate granice snage svakog generator, tj.

$$P_g^{\min} \le P_g \le P_g^{\max} \quad (5)$$

gde su: $P_g^{\min}$, $P_g^{\max}$ i P_g su minimalna, maksimalna i stvarna snaga g-tog generatora, i za zadatu ravnotežu između proizvedene snage i utrošene snage, tj.

$$\sum_{g\in G} P_g - P_D - P_{loss} = 0, \quad (6)$$

gde su: P_D ukupna snaga svih potrošača i P_{loss} gubici snage u prenosnom sistemu.

Gubici snage u prenosnom sistemu, P_{loss} , se izražavaju kao kvadratna funkcija trenutne snage generatora, tj. iz Kronove formule gubitaka [8], kao:

$$P_{loss} = \sum_{g\in G}\sum_{j\in G} P_g B_{gj} P_j + \sum_{g\in G} B_{0g} P_g + B_{00} \quad (7)$$

gde su B_{gj} i B_{0g} koeficijenti B-loss matrice a B_{00} je konstanta.

Da bi se zadovoljilo ograničenje (6), tokom iterativnog procesa optimizacije, jedan od generatora (npr. generator G) je odabran kao zavisni (labav) generator. Za taj generator vrednost izlazne snage, P_G, se računa iz sledeće jednačine:

$$P_G = P_D + P_{loss} - \sum_{g=1}^{G-1} P_g \quad (8)$$

Gubici snage, P_{loss}, se onda dobijaju na sledeći način: (i) zadavanje početne vrednosti $P_{loss} = P_{loss}^{(0)} = 0$ u (8), (ii) određivanje vrednosti $P_G^{(0)}$ iz (8) za $P_{loss} = P_{loss}^{(0)} = 0$, (iii) izračunavanje nove vrednosti $P_{loss}^{(1)}$ primenom (7), (iv) proveravanje da li je razlika između dve uzastopne vrednosti gubitaka snage manja ili jednaka zadatoj dozvoljenoj oleranciji δ, tj.

$$\left|P_{loss}^{(1)} - P_{loss}^{(0)}\right| \le \delta \quad (9)$$

i (v) izračunavanje vrednosti $P_G^{(1)}$ primenom (8) za $P_{loss} = P_{loss}^{(1)}$. Ako je razlika $\left|P_{loss}^{(1)} - P_{loss}^{(0)}\right|$ manja ili jednaka zadatoj toleranciji δ, ograničenje (6) koje predstavlja ravnotežu snaga, je zadovoljeno. U suprotnom, procedura se ponavlja. Kada je vrednost P_G izračunata, potrebno je proveriti da li se vrednost P_G nalazi u odgovarajućim granicama (5). Zatim se definiše promenljiva, $P_G^{\lim}$, na sledeći način:

$$P_G^{\lim} = \begin{cases} P_G^{\max} & ako\ \ je \quad P_G > P_G^{\max} \\ P_G^{\min} & ako\ \ je \quad P_G < P_G^{\min} \\ P_G & ako\ \ je \quad P_G^{\min} \le P_G \le P_G^{\max} \end{cases} \quad (10)$$

Da bi se osiguralo da zavisna promenljiva P_G ostaje u zadatim granicama, funkciji cilja (4) se dodaje kvadratni penalni član sa penalnim faktorom, λ_p. Na taj način se dobija proširena funkcija cilja:

$$FE_p = FE + \lambda_p\left(P_G - P_G^{\lim}\right)^2 \quad (11)$$

III. PSO I PPSO

Optimizacija roja čestica (eng. Particle swarm optimization (PSO)) je inspirisana ponašanjem rojeva u prirodi u potrazi za hranom [9]. Jedinka u jatu menja svoju poziciju i brzinu kretanja postepeno se krećući ka izvoru hrane. U PSO, svaka jedinka (čestica) u jatu je predstavljena vektorima pozicije i brzine, na sledeći način:

$$X_i(t) = \left[x_i^1(t),\ldots,x_i^k(t),\ldots,x_i^n(t)\right] \quad (12)$$

$$V_i(t) = \left[v_i^1(t),\ldots,v_i^k(t),\ldots,v_i^n(t)\right] \quad (13)$$

gde su $X_i\,(t)$ i $V_i\,(t)$ vektor pozicije i vektor brzine i-te čestice u vremenu (iteraciji) t; $x_i^k(t)$ i $v_i^k(t)$ su pozicija i brzina i-te čestice k-te dimenzije. Početne vrednosti vektora su slučajno odabrane. Brzine i pozicije čestica u narednoj iteraciji su određene pomoću sledećih jednačina:

$$\begin{aligned} v_i^k(t+1) &= w(t)v_i^k(t) + C_1 r_1\left(pbest_i^k(t) - x_i^k(t)\right) + \\ &+ C_2 r_2\left(gbest^k(t) - x_i^k(t)\right) \end{aligned} \quad (14)$$

$$x_i^k(t+1) = x_i^k(t) + v_i^k(t+1) \quad (15)$$

U (14), $w(t)$ je inercijalna težina, C_1 i C_2 su parametri regulacije ubrzanja čestica, r_1 i r_2 su are the uniformno raspoređeni slučajni brojevi unutar granica [0,1], $pbest_i^k(t)$ je najbolja pozicija i-te čestice k-te dimenzije (indiviualna najbolja pozicija), i $gbest^k$ je globalno najbolja pozicija u celoj populaciji. Drugi član u (14) predstavlja exploracioni deo PSO. Inercijalna težina vrši uravnoteženje između lokalnog i globalnog pretraživanja rešenja. U početnom stadijumu procesa pretraživanja vrednost w je velika kako bi se pojačala globalna eksploracija. U poslednjem stadijumu vrednost w se smanjuje kako bi se dobila bolja lokalna exploracija.

Algoritam PPSO je predložio Gholamghasemi M. sa koautorima 2019. godine [4]. Parametri C_1 i C_2, koji se u algoritmu PSO zadaju ručno, u PPSO algoritmu su modelovani faznim uglom (θ) definisanim u teoriji fazora. Na taj način, PPSO, za razliku od PSO, postaje adaptivni i neparametarski algoritam. Vrednost $w(t)$ je u PPSO jednaka nuli. Brzina u svakoj iteraciji se ažurira na sledeći način.

$$V_i(t) = \left|\cos\theta_i(t)\right|^{2\cdot\sin\theta_i(t)} \times \left(Pbest_i(t) - X_i(t)\right) + \\ + \left|\sin\theta_i(t)\right|^{2\cdot\cos\theta_i(t)} \times \left(Gbest(t) - X_i(t)\right) \tag{16}$$

gde su: $Pbest_i(t)$ i $Gbest(t)$ vektori individualne i globalne najbolje pozicije; $X_i(t)$ je vektor trenutne pozicije i-te čestice u t-toj iteraciji; θ_i jedno-dimenzioni fazni ugao vektora $\vec{X}_i\angle\theta_i$ za i-tu česticu. Za početnu populaciju koja se sastoji od N čestica (za t = 1), vektor $\vec{X}_i$ je: $\vec{X}_i = |X_i|\angle\theta_i$ (i = 1:N). Na početku pretraživanja rešenja, generisano je N slučajnih čestica (rešenja) u n-dimenzionom prostoru problema sa faznim uglom θ_i dobijenim iz ravnomerne raspodele $\theta_i = U(0, 2\pi)$, i sa početnom granicom brzine $V_{i,max}$. Donja i gornja granica $V_i(t)$ su definisane sledećim intervalom [$-V_{i,max}(t)$, $V_{i,max}(t)$].

Pozicije čestica se ažuriraju pomoću sledeće jednačine:

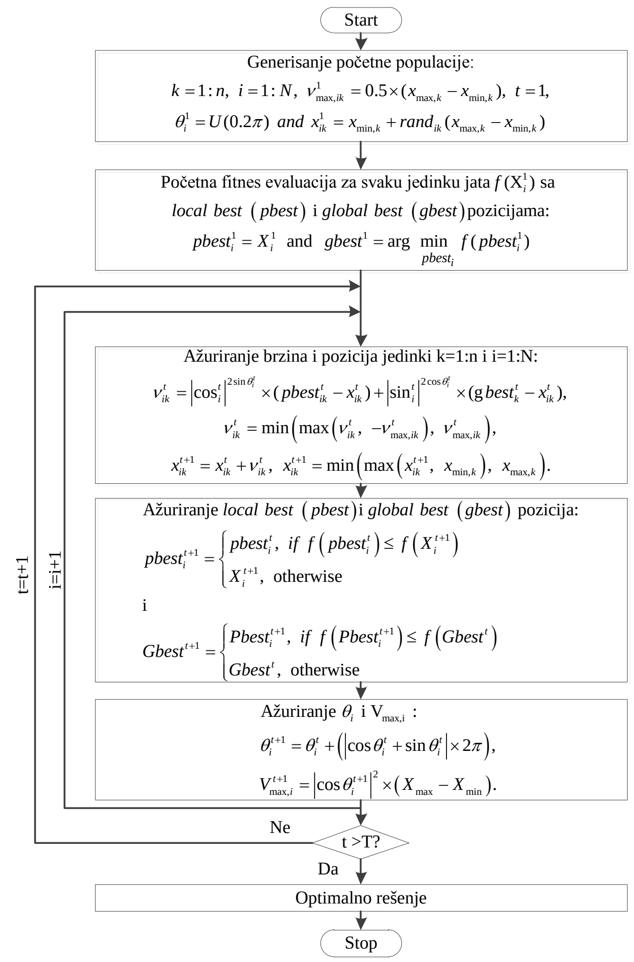
$$\vec{X}_i(t+1) = \vec{X}_i(t) + \vec{V}_i(t) \tag{17}$$

Posle ažuriranja brzine čestice i pozicije primenom (16) i (17), fazni ugao θ_i i maksimalna brzina $V_{i,max}$ za sledeću iteraciju izračunavaju se iz sledećih jednačina:

$$\theta_i(t+1) = \theta_i(t) + \left|\cos\theta_i(t) + \sin\theta_i(t)\right| \times (2\pi) \tag{18}$$

$$V_{i,\max}(t+1) = \left|\cos\theta_i(t)\right|^2 \times \left(X_{\max} - X_{\min}\right) \tag{19}$$

Na Sl. 1 nacrtan je dijagram toka PPSO.



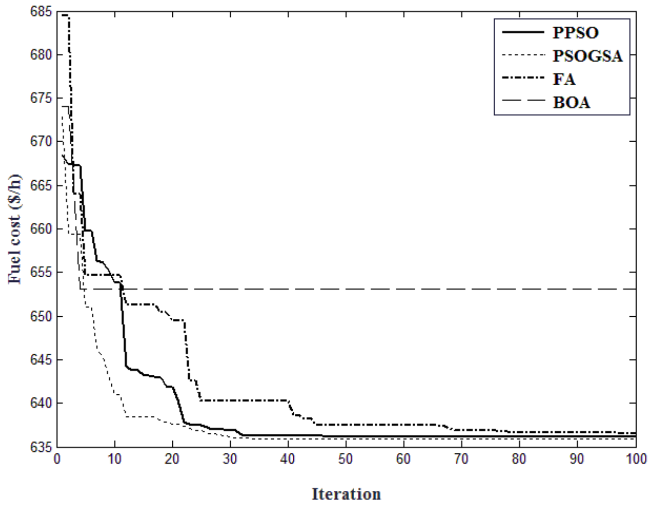
Sl. 1. Dijagram toka PPSO

IV. REZULTATI SIMULACIJE

Testiranje PPSO algoritma u ovom radu se vrši na standardnom IEEE test sistemu sa 30 čvorova, 6 generatora i ukupnom potrošnjom od 283.4 MW. Uzimaju se u obzir efekat ventila u termoelektranama i gubici snage u sistemu. *B-loss* matrica i koeficijenti troškova i emisije usvojeni su iz [8]. Implementacija PPSO se sprovodi na platformi od 1.6 GHz sa 3 GB RAM primenom MATLAB R2017a. Kao rezultati uzimaju se najbolje vrednosti dobijene posle 30 puštanja algoritma. Veličina dozvoljene greške u (9) je $\delta = 10^{-6}$ MW, dok je faktor skaliranja γ_{NOx} jednak 1,000 (\$/t). Minimizacija se vrši sa tri vrednosti težinskog faktora: w = 1 (minimizacija samo troškova goriva), w = 0 (minimizacija samo NO_x emisije) i w = 0.5 (istovremena minimizacija troškova goriva i emisije NO_x gasova). Rezultati dobijeni primenom PPSO upoređuju se sa rezultatima dobijenim pomoću tri sledeća algoritma: (i) hibridnog algoritma koji se sastoji od PSO i gravitacionog pretraživačkog algoritma (eng. PSO - Gravitational Search Algorithm (PSOGSA)) [10], koji je pokazao najbolje rezultate pri rešavanju CEED problema bez uzimanja u obzir efekta ventila [1], [11]; (ii) algoritma optimizacije leptira (eng. Butterfly Optimization Algorithm (BOA)) [12], kao jednog od najnovijih meta-heurističkih algoritama; i (iii) algoritma svica (eng. Firefly Algorithm (FA)) [13], kao jednog od najpoznatijih algoritama.

Konstante testiranih algoritama, koji se primenjuju u simulaciji, date su u Tabeli 1. U Tabeli 2 date su minimalne i maksimalne vrednosti rezultata i njihove standardne devijacije za primenjene algoritme. Iz Tabele 2 sledi da je minimalna vrednost troškova goriva, dobijena primenom PPSO najmanja u odnosu na minimalne vrednosti dobijene pomoću drugih testiranih algoritama. Minimalne vrednosti emisije NO_x gasova su iste u slučaju primene PPSO, PSOGSA i FA. Te vrednosti su bolje (manje) nego u slučaju primene BOA. Standardne devijacije rezultata dobijenih pomoću PPSO su manje nego standardne devijacije rezultata dobijenih pomoću PSOGSA, FA i BOA. U Tabeli 3 date su najbolje vrednosti izlaznih snaga generatora, troškova goriva i emisije gasova, dobijene primenom PPSO za $w = 1$, $w = 0$ i $w = 0.5$.

Na Sl. 3 date su krive konvergencije algoritama PPSO, PSOGSA, FA i BOA algorithms u slučaju minimizacije troškova goriva. Sa Sl. 3 se vidi da PPSO konvergira ka minimalnoj vrednosti za broj iteracija koji je isti kao u slučaju PSOGSA. U poređenju sa FA, PPSO konvergira ka minimalnoj vrednosti za manji broj iteracija. Broj iteracija BOA je manji u odnosu na ostale algoritme ali BOA daje lošije vrednosti minimalnih troškova goriva, emisije gasova i standardne devijacije rezultata. Sl. 3 pokazuje da su početne brzine konvergencije velike za sve primenjene algoritme.



Sl. 3. Krive konvergencije PPSO, PSOGSA, FA i BOA u za slučaj minimizacije troškova goriva.

TABELA 1
KOEFICIJENTI ALGORITAMA KOJI SU TESTIRANI NA STANDARDNOM IEEE TEST SISTEMU SA 30 ČVOROVA I 6 GENERATORA

PPSO		PSOGSA						FA					BOA				
N	T	N	T	G_0	α	C_1	C_2	N	T	A	β_{min}	γ	N	T	c	a	p
50	200	50	200	1	20	0.5	1.5	50	200	0.25	0.2	1	50	200	0.01	0.1	0.8

TABELA 2
MINIMALNE I MAKSIMALNE VREDNOSTI I STANDARDNE DEVIJACIJE, DOBIJENE PRIMENOM PPSO, PSOGSA, FA I BOA NA STANDARDNOM IEEE TEST SISTEMU SA 30 ČVOROVA I 6 GENERATORA

Algoritam		PPSO	PSOGSA	FA	BOA
Minimizacija troškova goriva ($w = 1$)	Min	635.82129	635.82284	635.83288	640.37240
	Max	647.29186	698.99430	642.65875	663.92341
	SD*	2.376452	18.37740	2.904691	5.989508
Minimizacija emisije NO_x gasova ($w = 0$)	Min	0.1941785	0.1941785	0.1941785	0.1942077
	Max	0.1941785	0.2195708	0.1941785	0.1966057
	SD*	5.8637e-11	6.23630	1.0606e-10	5.7676e-04
$w = 0.5$	SD*	2.9438e-2	9.68220	1.96486e-1	2.5305474

* SD označava standardnu devijaciju

V. ZAKLJUČAK

U ovom radu je predložen algoritam PPSO za rešavanje CEED problema. Performanse ovog algoritma pri rešavanju CEED problema su procenjivane na standardnom IEEE test sistemu sa 30 čvorova i 6 generatora. Pri tome, uzimani su u obzir uticaj efekta ventila u termoelektranama i gubici snage u elektroenergetskom sistemu. Zatim su dobijeni rezultati upoređeni sa rezultatima drugih algoritama: PSOGSA koji je u radu [1] pokazao najbolje rezultate pri rešavanju CEED problema na IEEE test sistemu sa 30 čvorova i 6 generatora ali bez uzimanja u obzir efekta ventila; BOA, koji predstavlja jedan od najnovijih meta-heurističkih algoritama; FA, koji je jedan od često primenjivanih algoritama. Poređenjem testiranih algoritama, utvrđeno je da PPSO daje najbolje rezultate: Simulacioni rezultati su pokazali da PPSO ima dobre konvergentne osobine i daje najbolje vrednosti minimalnih troškova goriva u odnosu na algoritme PSOGSA, FA i BOA. Osim toga, utvrđeno je da su standardne devijacije rezultata najmanje u slučaju primene PPSO, da su minimalne vrednosti emisije štetnih gasova iste u slučajevima primene PPSO, PSOGSA i FA i da su one bolje nego u slučaju primene BOA.

TABELA 3
NAJBOLJE VREDNOSTI IZLAZNIH SNAGA, TROŠKOVA GORIVA I EMISIJE GASOVA, DOBIJENE PRIMENOM PPSO

Snaga, MW	w = 1	w = 0	w = 0.5
$P_{s,1}$	5.00000	41.09207	5.00000
$P_{s,2}$	13.44427	46.36641	18.32689
$P_{s,3}$	83.53982	54.44192	79.88927
$P_{s,4}$	74.84721	39.03759	74.81317
$P_{s,5}$	79.79982	54.44609	78.55621
$P_{s,6}$	28.65457	51.54889	28.76874
P_{loss}	1.88568	3.53297	1.95428
Troškovi goriva ($/h)	635.82129	728.66678	638.65784
NO_x (ton/h)	0.226433	0.1941785	0.223048

DODATAK

TABELA 4
B-LOSS MATRICE TEST SISTEMA [8]

Mat-rice	Elementi matrica
B	$\begin{bmatrix} 0.1382 & -0.0299 & 0.0044 & -0.0022 & -0.0010 & -0.0008 \\ -0.0299 & 0.0487 & -0.0025 & 0.0004 & 0.0016 & 0.0041 \\ 0.0044 & -0.0025 & 0.0182 & -0.0070 & -0.0066 & -0.0066 \\ -0.0022 & 0.0004 & -0.0070 & 0.0137 & 0.0050 & 0.0033 \\ -0.0010 & 0.0016 & -0.0066 & 0.0050 & 0.0109 & 0.0005 \\ -0.0008 & 0.0041 & -0.0066 & 0.0033 & 0.0005 & 0.0244 \end{bmatrix}$
B_0	$\begin{bmatrix} -0.0107 & 0.0060 & -0.0017 & 0.0009 & 0.0002 & 0.0030 \end{bmatrix}$
B_{00}	$\begin{bmatrix} 0.00098573 \end{bmatrix}$

TABELA 5
KOEFICIJENTI TROŠKOVA GORIVA I EMISIJE NO_X GASOVA I OGRANIČENJA GENERATORA ZA PRIMENJENI TEST SISTEM [8]

g	a_g	b_g	c_g	d_g	e_g	α_g	β_g	η_g	ξ_g	λ_g	P_g^{min}	P_g^{max}
1	10	200	100	18	3.7	4.091e-2	-5.554e-2	6.490e-2	2.0e-4	2.857	5	150
2	10	150	120	16	3.8	2.543e-2	-6.047e-2	5.638e-2	5.0e-4	3.333	5	150
3	20	180	40	14	4.0	4.258e-2	-5.094e-2	4.586e-2	1.0e-6	8.0	5	150
4	10	100	60	12	4.5	5.326e-2	-3.550e-2	3.380e-2	2.0e-3	2.0	5	150
5	20	180	40	13	4.2	4.258e-2	-5.094e-2	4.586e-2	1.0e-6	8.0	5	150
6	10	150	100	13.5	4.1	6.131e-2	-5.555e-2	5.151e-2	1.0e-5	6.667	5	150

ZAHVALNICA

Ovaj rad je finansijski pomognut od strane Ministarsva prosvete, nauke i tehnološkog razvoja Republike Srbije.

LITERATURA

[1] J. Radosavljević, "A solution to the combined economic and emission dispatch using hybrid PSOGSA algorithm," *Appl. Artif. Intell.* vol. *30*, no. *5*, pp. *445-474*, Jun. 2016.

[2] D. W. Gong, Y. Zhang, C. L. Qi, "Environmental/economic power dispatch using a hybrid multi-objective optimization algorithm," *Int. J. Elec. Power*, vol. *32*, pp. *607-614*, Nov. 2010.

[3] M. Jevtić, N. Jovanović, J. Radosavljivić, D. Klimenta, "Moth swarm algorithm for solving combined economic and emission dispatch problem," *Elektron. Elektrotech.* vol. *23*, pp. *21-28*, Jun. 2017.

[4] M. Gholamghasemi, E. Akbari, A. Rahimnejad, S. E. Razavi, S. Ghavidel, L. Li, "Phasor particle swarm optimization: a simple and efficient variant of PSO," *Soft Comput.* vol. *23*, no. *19*, pp. *9701-9718*, March 2019.

[5] L. Benasla, A. Belmadani, M. Rahli, "Spiral optimization algorithm for solving combined economic and emission dispatch," *Int. J. Elec. Power.* vol. *62*, pp. *163-174*, Nov. 2014.

[6] A. Bhattacharya, P. K. Chattopadhyay, "Solving economic emission load dispatch problems using hybrid differential evolution," *Appl. Soft Comput.* vol. *11*, no. *2*, pp. *2526–2537*, March, 2011.

[7] U. Güvenç, Y. Sönmez, S. Duman, N. Yörükeren, "Combined economic and emission dispatch solution using gravitational search algorithm," *Sci. Iran.* vol. *19*, no. *6*, pp. *1754–1762*, Dec. 2012.

[8] D. Aydin, S. Ozyon, C. Yasar, and T. Liao, "Artificial bee colony algorithm with dynamic population size to combined economic and emission dispatch problem," *Int. J. Elec. Power*, vol. *54*, pp. *144–153*, Jan. 2014.

[9] J. Kennedy, R. Eberhart, "Particle swarm optimization**,"** Proc. ICNN95' International Conference on Neural Networks, IEEE, Perth, WA, Australia, pp. *1942–1948*, 27 Nov. - 1 Dec. 1995.

[10] S. Mirjalli, S. Z. M. Hashim, H. M. Sardroudi, "Training feedforward neural networks using hybrid particle swarm and gravitational search algorithm," *App. Math. Comput*. vol. *218*, no. *22*, pp. *11125-11137*, Jul. 2012.

[11] J. Radosavljević, *Metaheuristic optimization in power engineering*, London, United Kingdom: IET, 2018.

[12] S. Arora, S. Singh, "A hybrid optimization algorithm based on butterfly optimization algorithm and differential evolution," *Int. J. Swarm Intelligence*, vol. *3*, no. **2/3**, pp. *152-169*, Oct. 2017.

[13] X. S. Yang, "Firefly algorithm, L´evy flights and global optimization," in *Research and development in intelligent systems XXVI*, vol. *26*, pp. *209-218*, London, United Kingdom: Springer-Verlag, 2010.

ABSTRACT

Minimization of fuel costs and pollutant emissions in thermal power plants by adjusting electric power outputs from generators represents important problem in power system management. This problem is known as Combined economic emission dispatch (CEED) problem. In this paper, a meta-heuristic algorithm called Phasor particle swarm optimization, which is an improved variant of Particle swarm optimization, is proposed to solve the CEED problem. Parameters of Phasor particle swarm optimization are automatically adjusted during iterations, so this algorithm is adaptive and non-parametric, which is its advantage. The performance of the proposed algorithm for solving CEED problem is evaluated in a standard IEEE test system with 30 nodes and 6 generators. Based on the obtained results, it was determined that this algorithm has better characteristics than the algorithms used in other published papers to solve the CEED problem.

Solving combined economic and emission dispatch problem using Phasor particle swarm optimization

Milena Jevtić, Miroljub Jevtić, Jordan Radosavljević, Sanela Arsić i Dardan Klimenta

Potential of Using Simulated Data in Processing Photoacoustic Measurement Data

M. I. Jordovic Pavlovic, A. D. Kupusinac, S. P. Galovic, D. D. Markushev, M.N. Nesic, K.Lj.Djordjevic, M. N. Popovic

Abstract: This paper explores the potential of using simulated data in calibration of photoacoustic measurement system. The database of simulated experimental values is created using software developed on the bases of the theory-mathematical model. Reliability of the data was gained thanks to the expert knowledge. An artificial neural network as a precise prediction tool is trained on the developed database of simulated data to recognize type of the microphone used as a detector in photoacoustic experiment. The result is classification model satisfies the basic requirements of a photoacoustic experiment: accuracy, reliability and real time operations. The paper discusses the optimization of classification model in terms of used computational power, required precision and process rate in relation with defined problem. The obtained results justify the idea of using simulated data in photoacoustic. Presented theory-mathematical model and classification model are part of developed machine learning framework for processing photoacoustic measurement data.

Keywords: Machine learning, artificial neural networks, simulated data, classification, photoacoustics, microphone

I. Introduction

Machine learning techniques are considered suitable tool for intelligent decision making, and therefore they have found application in various domains. When input and output parameters are linked with some kind of pattern, and sufficient data is available, this pattern can be discovered or approximated by machine learning algorithm being trained on that same data. Subsequently, output for particular inputs outside the learning dataset can be calculated (with more or less accuracy) using this newly discovered pattern. This means that if the quality and the quantity of the data used for learning are sufficient, and the discovered pattern also exists for events that were not part of the learning dataset, the produced result can be used to approximate the outputs based on any future input[1]. Machine learning algorithms, and, in particular, artificial neural networks (ANN), are frequently used as reliable and fast prediction tools. They are often used in photoacoustics (PA), a popular method in photothermal (PT) science in the last few years, for: noise removal in photoacoustic recognition of images [2], simultaneous determination of the laser beam spatial profile and relaxation time of the polyatomic molecules in gases in real time within the trace atmosphere gases monitoring [3][4], reconstruction of optical profile of optically gradient materials based on frequency, magnitude and phase of measured PT response [5], etc.

Miroslava Jordović Pavlović is with the Western Serbia Academy of Applied Studies, Užice, Trg Svetog Save 34, 31000 Užice, Serbia (miroslava.jordovic-pavlovic@vpts.edu.rs).
Aleksandar Kupusinac is with Faculty of Technical Sciences, University of Novi Sad, 6 Trg Dositeja Obradovića, 21000 Novi Sad, Serbia (sasak@uns.ac.rs).
Slobodanka Galović is with the Vinča Institute of Nuclear Sciences, 12-14 Mike Petrovića Alasa, 11351 Vinča, Serbia (bobagal@vin.bg.ac.rs).
Dragan Markushev is with the Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Pregrevica, Serbia (dragan.markushev@ipb.ac.rs).
Mioljub Nešić is with the Vinča Institute of Nuclear Sciences, 12-14 Mike Petrovića Alasa, 11351 Vinča, Serbia (mioljub@gmail.com).
Katarina Đorđević is with the Faculty of Physics, University of Belgrade Studentski trg 12, 11000 Beograd, Serbia (katarinaljdjordjevic@gmail.com).
Marica Popović is with the Vinča Institute of Nuclear Sciences, 12-14 Mike Petrovića Alasa, 11351 Vinča, Serbia (maricap@vin.bg.ac.rs)

In this paper, a few of the several results achieved in PA measurement system characterization research are presented. The ultimate goal is material characterization. The aim of the PA measurements is the determination of physical properties (thermal, optical, mechanical, elastic, electronic and other related ones) of the examined structure from its PA response. All PT methods are indirect measurement techniques, and so is the photoacoustics, meaning that these methods are model dependent. In terms of mathematics, obtaining physical properties by these methods is considered an inverse problem that can be assessed in two steps:

1. **Development of the direct (forward) model – direct solution of the inverse problem,** i.e. developing the mathematical model that sufficiently well describes physical processes leading from the optical excitation to the thermal response. First step is theoretical-mathematical modeling of temperature distribution within the sample, on front and back sample surfaces and in its surroundings, and then theoretical-mathematical modeling of the specific PT response (in this case the PA response)
2. **Development of the inverse procedure – inverse solution of the inverse problem,** i.e. the determination of physical properties of the sample based on measured photothermal response, developed mathematical model and well known preset of input parameters (the intensity and modulation frequency of the incident optical radiation). Some of the inverse procedures are fitting, numerical procedures and neural networks. Fitting and numerical procedures are time consuming procedures, demanding the engagement of the researcher. These are drawbacks regarding scientific and further industrial application of the method, where a real time procedure is appreciated. The reasonable choice is

artificial neural networks as a very efficient machine learning algorithm. Because of the complexity of the inverse problem more than one ANN is needed.

Firstly, the characterization of the sample and the prediction of its thermal, mechanical and optical properties based on its PA response, require the use of one ANN. This is already done in [6]. But the necessary precondition is that the PA response in use is influenced only by the sample, which, unfortunately, is not the case. The PA response is non- linearly affected not only by the sample but also by the measurement instrument chain and the appearing noise. So, the PA response has to be corrected first, in order to obtain the so called "true" signal. In the first preprocessing procedure that was developed, noises are removed during data acquisition. In the second preprocessing procedure, calibration of the measurement system has to be done. Because of the dominant impact of the microphone on the distortions in the measurement instrument chain, as the consequence of using minimum volume cell configuration of the PA experimental set-up, calibration of the measurement system boils down to the calibration of the microphone. The key of this brand new idea is the determination of microphone transfer function. Furthermore, the division of PA response amplitude data by corresponding microphone characteristics and its subtraction from the PA response phase data will result in gaining the so called "true" signal, originating only from the sample. Unfortunately, microphone specifications provided by the manufacturer are not precise enough, particularly in the case of phase transfer function. Besides, microphone cavity is not considered as the source of resonances, which is inevitable in PA measurements. Since, these specifications could not be used, the other solution is needed. Non-linear influence of the microphone on a PA response suggests ANN application. Having in mind that ANN seeks large datasets (is data hungry) [7], the first requirement for the application of neural networks is set. But this requirement is opposed to two facts related to PA measurements: firstly, such a numerous experimental collection is very difficult to obtain, and secondly, based on the experience, real experimental data can hide a very serious problem of the influence of the measurement system on the estimated parameter values [8]. Therefore, another solution for database creation is presented: the idea of theoretical-mathematical model as a base for designing a software for the simulation of PA experimental values. Thanks to the developed software, amplitude and phase data of the simulated PA response are obtained. Here, satisfactory credibility to the experiment is of essential importance in order to make the newly created method precise enough. Therefore, expert knowledge (i.e., the preset input parameters) is crucial for the solution of this problem.

Simulated data have often been used for training in machine learning problems in the past few years [9][10][11], but as far as we know, the idea of using simulated experimental values, obtained by developed software based on a theoretical - mathematical model, for training a machine learning model is new. This article presents a few steps of a complex correction procedure performed in photoacoustic measurements. Firstly, a complete method of making a large amount of reliable simulated data as a precondition for applying neural networks as the inverse solution of the inverse problem is explained. Secondly, a process of designing classification model for microphone type recognition as the first step in recognizing measurement system characteristics is discussed on the base of optimal computational complexity, required precision and process rate in relation to the given problem and available data set. Classification of microphone type will determine the shape of the transfer function and the levels of signal exaggeration and attenuation. Once the class of the microphone is defined, characterization of microphone will be simplified by limiting the database made for various types of microphones to a database of a particular microphone type. This idea is presented in our previous work [12]. That way, time, and computational power are saved, which are real benefits of the classification model. Learning on the defined database of classified microphone type, ANN based model for microphone characterization [13] predicts characteristic microphone parameters with satisfying accuracy, which together with the corresponding shape, precisely determine microphone transfer function [12].

This paper shows that if a massive dataset is obtained and the quality of data is high, less computation power is needed, and higher process rate is gained for the solution of machine learning problem.

II. Theoretical -Mathematical Model of Photoacoustic Response

Photoacoustics, as one of photothermal methods, is based on the photothermal effect. The photothermal effect is the effect of generation of heat as a consequence of the absorption of the incident electromagnetic radiation, from a wide spectrum of wavelengths, in different relaxation and de-excitation processes. This way generated heat causes the disruption of the thermodynamic state of the sample (pressure, temperature, density) which propagates through the sample and the nearby environment, producing a number of detectable phenomena. In photoacoustics, the first and the most used photothermal method, a sample is placed inside the photoacoustic cell that contains air and microphone. It is exposed to a modulated light beam which causes periodic sample heating. As a consequence, the air pressure in the PA cell oscillates, which can be detected by a microphone [14].The photoacoustic cell can be designed in a so called "reflection configuration", with the source and the microphone set up on the same side of a sample, or the "transmission configuration", where a sample is placed between the source and the microphone. In our experimental set up, the minimum volume cell configuration is employed. It is kind of transmission cell configuration where sample is mounted directly on the top of the microphone, instead of the dust cover, as presented in figure 1, [15]. This way, the microphone chamber acts as the PA cell, closed by the sample on one side and the microphone diaphragm on the other one, which causes disruptions of the recorded signal on its endings [16].

Power levels of experimentally recorded signals are, generally, low. In order to make the level of the recorded signal higher than the level of the noise (real – flicker noise, coherent signal deviation and random noise), the absorption of the sample has to be large. In the case of materials with significant reflection, the additional coating is needed, while in the case of transparent (or semitransparent) samples, the coefficient of transmission has to be augmented. However, due to high level of transmission, the absorption of the incident radiation in the surrounded air can`t be neglected and the recorded signal begins to contain unnecessary information. The problem is even bigger in the case of the minimum volume cell configuration, where the microphone has to be protected because of the small dimensions of the cell. Another solution is, also, an additional layer of high absorption.

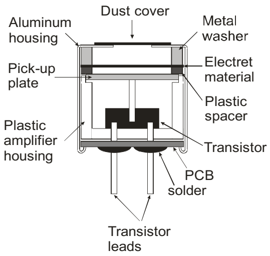


Fig. 1. Experimental setup

Photoacoustic response within the transmission configuration is the sum of two dominant signal components: thermoconducting and thermoelastic component. Thermoconducting component arises due to the periodic heat flow from the sample to the surrounding gas (thermal-piston effect) and thermoelastic component arises due to the thermoelastic banding of the sample (drum effect) [17][18][19][20][21][22][23][24][25].

In our experiments two-layer structure is employed. The first layer is black coating, and the second layer is the investigated sample. Theoretical-mathematical model of PA response of a two-layer system, used for obtaining the dataset, is given by following expressions [26][27]:

$$\tilde{p}_{tot} = \tilde{p}_{th} + \tilde{p}_{ac} \tag{1}$$

$$\tilde{p}_{th} = \delta P = \frac{\gamma P_0}{l_g T_0} \frac{1}{\tilde{\sigma}_g} \tilde{\vartheta}(l_s) \tag{2}$$

$$\tilde{p}_{ac} = \frac{3\gamma P_0}{l_a} \frac{R^2}{l_s^3} \left[\alpha_{T1} \int_0^{l_1} \left(x - \frac{l_s}{2}\right) \tilde{\vartheta}(x) dx + \alpha_{T2} \int_{l_1}^{l_s} \left((x - l_1) - \frac{l_s}{2}\right) \tilde{\vartheta}(x) dx \right] \tag{3}$$

Where p_{tot} is total pressure that we want to record by photoacoustic, p_{th} is the thermoconducting component and p_{ac} is the thermoelastic component. Furthermore P_0 is the presser in the cell, V_0 is the volume of the cell (in the case of the minimum volume cell V_0 represents the volume of the chamber cavity), γ represents the heat capacity ratio, α_T is the thermal expansion coefficient, R_c is the radius of the chamber in front of the microphone diaphragm, $l_1$and l_2 are the thicknesses of the first and second layer, while l_s is the sum of the thicknesses these two layers ($l_1$and l_2). $\vartheta(x)$ represents temperature variations inside the samples and $\vartheta(l_s)$ is the surface temperature variation on the rare surface. Expressions for these temperature variations are given in the article [28][29]. The presented model described the total presser as photoacoustic response, and its components in the two-layer system surrounded by the air and it is based on the Generalized model of heat conduction that implies finite heat propagation speed. The system depicts volume absorption of incident optical beam in both layers [26][27][28][29].

Appearance of amplitude and phase characteristics of the theory-mathematical simulated total pressure are shown on figure 2a) and 2b) respectively.

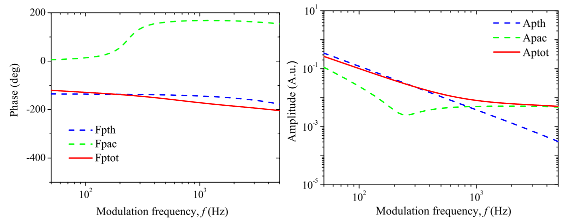


Fig. 2 Simulated amplitude and phase (solid line) of the total photoacoustic signal, $p_{tot}(f)$, as a function of the modulation frequency f, together with the appropriate components $p_{th}(f)$ and $p_{ac}(f)$ (dotted lines).

In a minimum volume cell PA experiment [30], microphone is the fundamental part of the detector system. Microphone is an acoustic-electric converter, but its transfer function in frequency and time domain differ due its construction, applied geometry and membrane type. In the literature [8] and in our experimental experience, microphone behavior is described as filtering. At low frequencies (< 1 kHz), electret microphones (commonly used in PA) usually act as electronic high-pass filters, while at high frequencies (> 1 kHz) these microphones usually act as acoustic low-pass filters.

The influence of the measurement chain, including the microphone as the component that has the greatest impact in signal distortion, is given by the following mathematical expressions describing total transfer function:

$$H_1^e(f) = \frac{1}{1 - j\frac{f_1}{f}} \tag{4}$$

$$H_2^e(f) = \frac{1}{1 - j\frac{f_2}{f}} \tag{5}$$

$$H^a(f) = \frac{f_3^2}{f_3^2 - f^2 + jff_3\xi_3} + \frac{f_4^2}{f_4^2 - f^2 + jff_4\xi_4} \tag{6}$$

$$H_{mic}(f) = H_2^e(f) H_{total}^a(f) \tag{7}$$

$$H_{total}(f) = H_1^e(f) H_{mic}(f) \tag{8}$$

In previous equations, $H_1^e(f)$ represents electronic characteristic of the influence of the other components in the measurement chain, first of all the sound card, and f_1 is the characteristic frequency that describes this system. Based on experimental experience, it is assumed that this frequency is constant. $H_2^e(f)$and $H^a(f)$ represent electronic and acoustic characteristics of the microphone. f_2 correspondes to the characteristic frequency of the electronic high-pass filter and f_3 and f_4 to the characteristic frequencies of the acoustic low-pass filters of the microphone, ξ_3 and ξ_4 are reciprocial values of the quality factor, or, in other words, the double value of the damping factor. The product of these two components represents the microphone response. As a consequence, the microphone response in frequency domain is deviated in amplitude and phase, especially at the begging and at the end of frequency range. Different microphone types have different transfer functions, but transfer functions of two microphones of the same type are usually different, because, in practice, two identical microphones do not exist. Theoretical-mathematical model for the total photoacoustic signal recorded by the minimum volume cell photoacoustic experimental set up represents product of the total pressure and the total transfer function:

$$S(f) = \sigma p_{total}(f) H_{total}(f) \tag{9}$$

Based on this equation and numerical simulations of the experiments, the database is obtained. Amplitude and phase data of the simulated PA response are given in figure 3a) and 3b). All the curves (amplitude and phase) of distorted photoacoustic signal have expected shape, according to experimental experience. There are no outliers.

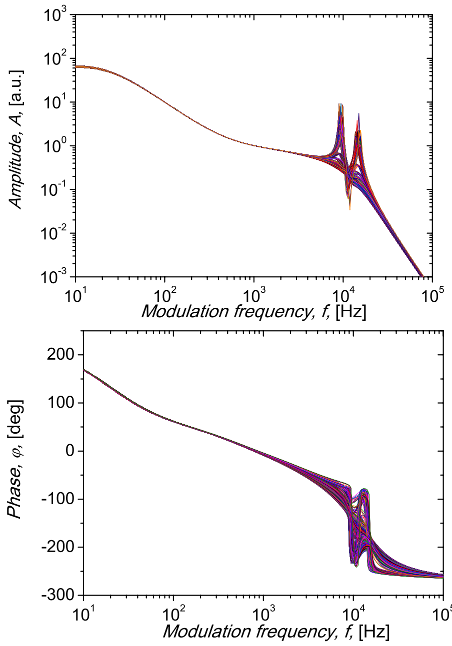


Fig. 3 Curves a) amplitude and b) phase of distorted photoacoustic signals with different microphone characteristics from the dataset used for network training[12]

III. Database description

Based on theoretical-mathematical model, software for creating simulated experimental values or numerical experiments is designed using programming IDE of Matlab. Microphone theoretical characteristics, corresponding to commercially available microphones ECM30B, ECM60 and WM66, are given in Table 1. Beside these microphone types, frequently used in PA experiments, simulations for another type of microphone are created, the so called ideal microphone (IM). Considering ideal microphone is of great importance for the correction procedure. If a microphone exerted ideal behavior, meaning it had flat PA response, that would mean that measurement chain would be equally sensitive in the whole frequency domain, so the correction procedure would be unnecessary. So, taking IM into account, we are saving the time.

TABLE I

Theoretical values for aluminum sample

	Dye	Aluminum
Thermal conductivity [$Wm^{-1}K^{-1}$]	70	210
Thermal diffusivity [m^2s^{-1}]	$2.5*10^{-5}$	$8.6*10^{-5}$
Thermal relaxation time [s]	10^{-4}	10^{-12}
Absorption coefficient [m^{-1}]	10^{8}	$145*10^{6}$

During the process of the examination, the black dye-aluminum structure was investigated. Aluminum plate, 197 µm in thicknesses and with radius of 10 µm was covered in black ink dye, 2 µm in thicknesses. Thermal, thermal memory and optical parameters used for obtaining database are given in Table 1.

Expert knowledge was crucial in obtaining similarity good enough with the experiment. Based on experimental experience, characteristic microphone parameters are considered to have different stability, regarding the reproducibility in each measurement. Accordingly, different value ranges were set for different parameters. Frequency f_2 is the most stable parameter due to its origin from RC microphone characteristic, so three values ware taken for network training: central value $f_{20} = 25$ Hz for the microphone ECM30B and two values which are ± 5 % apart from the central value (23.75 Hz and 26.25 Hz). By analogy, the values for the ECM60 are: 14.25Hz, 15Hz and 15.75 Hz, for the WM66 they are: 61.75 Hz, 65 Hz and 68.25 Hz, while for IM the values are: 0.475Hz, 0.5 Hz and 0.525 Hz. Frequencies f_3 and f_4 are more dependent on experimental conditions then f_2, so they are less stable than f_2. Ten values, equally distanced in the corresponding ranges, were considered to be good enough for the description of experimental behavior related to those two frequencies. f_{30} is taken in the range 8930-9866 Hz and f_{40} is taken in the range 13965-15432 Hz for ECM30B. Microphones ECM60 and WM66 have the same ranges for frequencies f_3 and f_4, 7980-8817 Hz and 13015-14383 Hz respectively. For IM f_3 is in the range of 190000-209998 Hz while f_4 is in the range of 285000-314997 Hz respectively. Damping factors of the second order low-pass filter ξ_3 and ξ_4 are strongly dependent on experimental conditions and they are the most unstable parameters. Each value range, for ξ_3 and ξ_4, was chosen based on the peak appearing in the amplitude characteristic of the second order filter. Critical value of quality factor in the case of limitary situation where signal is extremely damped and respectively unforced is Q=0.5. Significant change happens from Q=1 to Q=100 hence $\xi \in [0.99, 0.015]$. 15 values, irregularly distributed in this range, were taken for the each type of microphone. This kind of microphone parameter distribution was assumed to be good enough to simulate all possible experimental situations. The discussion and comparison of inverse problem-solving concepts in photoacoustics is presented in our previous work [31]. There are 65,000 paired curves for each microphone type, as 65,000 simulated experimental results, and those are 65,000 records of the database. Paired curves (two curves) mean that there are both amplitude and phase data for the given set of microphone parameters. Each curve contains data sampled at 200 frequency values in the range from 10Hz to 100kHz. By taking such a wide frequency domain, the possibility of using microphones with different membrane material (mylar, nickel, graphene) is considered. In total, every record is represented with 400 samples, 200 samples of amplitude and 200 samples of phase characteristics. Those are features for our machine learning problem. In other words each frequency is presented with two features, sample of amplitude and sample of phase, so we have resolution of two for every point on frequency axes. At the end of each database record, the information about which microphone type a particular record belongs to is written. The classification problem has 4 classes of microphones, symbolically presented with 0, 1, 2 and 3.

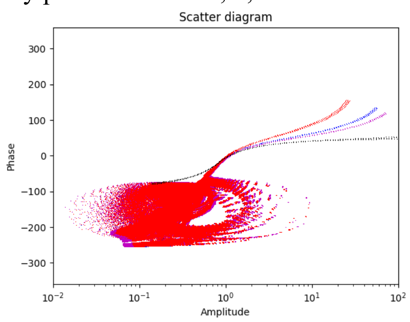


Fig.4. Visualization of the data, different colors correspond to different microphone types

Visualization of the data used in classification modeling, the form of scatter diagram, is given in Fig. 2. Each point on a scatter diagram is one point of 200 points that corresponds to one curve of 270,000 curves in the database. Different classes of microphone are presented with different colors. Analyzing the diagram, one can conclude that points are completely classified to four classes or four microphone types in upper-right part of the diagram, meaning for certain distribution of amplitude and phase values it is clear to which class point belongs. That distribution of amplitude and phase values are happening in a low frequency domain. In lower-left part of the diagram points are mixed, meaning that for that distribution of amplitude and phase values it is not clear to which class point belongs, i.e. curves (or classes) overlap. Thus, classification model has more difficult task because of the overlap. Training, validation and test sets are obtained randomly because dataset is first shuffled and then divided into training, validation and test set. Generalization of the results is obtained on that way, thus 243 000 records or 90% of the total number of records belongs to the training set, 13500 records or 5% belong to the validation set and the rest belongs to the test set.

IV. Research results and discussion

Once, the topology of the model is chosen, the next step is fine-tuning of topology itself, parameters and hyperparameters of the model. It is done in iterative process idea-code-experiment, with a numerous attempt using literature suggestions [32][33] and experience.

In pre-processing step, data scaling was done by performing the normalization of the input and output. Max normalization was chosen. It means that each element x_i of the input vector is divided by its maximum absolute value, which is the maximum of absolute values of all the samples, a total of 270000 values, at the i-th frequency. In other words, it is absolute maximum value of the i-th row of the input matrix. This way normalization of the input vector is done, all the values of the input vector are equal or less than unity. Similarly, normalization of the output vector is done. For weights parameters initialization, among others Xaviar algorithm [34] is chosen. The activation function *tanh()* is used for forward propagation and the Adam algorithm [35] is used for the optimization of weights in backpropagation. The optimization is intensified by the Mini-batch technique, size of 128. Because of the classification function softmax in the last layer, a cross entropy with logits is used as the error function and system performance measure during training. Neural network tuning on number of hidden layers and the number of neurons is presented in Table 2.

TABLE II

NUMBER OF HIDDEN LAYERS AND NUMBER OF NEURONS IN HIDDEN LAYER(S) ANALYSES

No.of hidden layers	1	2	2	2	1	2	2	1	2	2	1	2	1
No. of neurons of the 1. h. l.	10	8	7	9	5	3	4	4	2	3	3	2	2
No. of neurons of the 2. h. l.	/	2	3	1	/	2	1	/	2	1	0	1	0
Train accuracy(%)	99.99	99.99	99.99	75.02	99.99	99.99	50.03	99.99	99.99	81.15	99.99	75.01	99.99
Dev accuracy(%)	99.99	99.99	99.99	74.15	99.99	99.99	49.19	99.99	99.99	82.09	99.99	75.03	99.99
Test accuracy(%)	99.99	99.99	99.99	75.45	99.99	99.99	50.15	99.99	99.99	81.11	99.99	74.7	99.99
Number of epochs	100	100	100	100	100	100	100	100	100	100	100	100	100
Prediction time (ms)	14.34	17.89	17.44	/	14.06	16.75	/	13.89	16.89	/	13.73	/	13.93

According to Table 2, for the defined classification problem and the dataset of 270000 records following conclusions can be drawn. One neuron in second hidden layer in configuration of two hidden layers is not appropriate and those topologies were dismissed, but 2 neurons in second hidden layer are satisfying. The reason are 4 classes at the output. There is no difference in accuracy in the case of the configuration with one hidden layer and in the case of configuration with two hidden layers with same total number of neurons. Based on experimental experience one can say that for other machine learning problems that was not a case. This is specificity of this particular problem. So, the topology and the choice of model parameters and hyperparameters are singularity of machine learning problem and the quantity and quality of available data. Minimum configuration that satisfies required accuracy is one hidden layer with 2 neurons. It is surprisingly small number of neurons, which can be justified with the large data set. It means that learning with large datasets decreases the number of computational units of ANN configuration, it becomes computationally simpler. Large dataset brings into the model huge knowledge about the problem, in the case of our classification problem knowledge about photoacoustic experiment environment. Using this knowledge ANN needs less computational power and less epochs for learning. Analyzing the obtained prediction time of different topologies of classification model, the most important influence on the processing rate has the number of hidden layers, the number of neurons in layer has minor influence, even if there is significant difference in the number of neurons.

Concerning the prediction, the network gives very high accuracy, train, dev and test accuracy are equal, 99.99%. Concerning the training, the network obtained good results even for very quick training, that lasts 100 epochs. According to the equal values of training, dev and test accuracy and low error function on the new data sets we can conclude that the network generalizes very well. There is no overfitting.

The reliability of the model was tested on simulated data. Sixteen different independent datasets, meaning four different amplitude and phase characteristics for each type of microphone were created, where the microphone parameter values differed from those on which the network was trained, but in the given parameter range. Results are presented in Table 3. According to Table 3 our model is reliable, it recognizes the microphone type precisely and gives an answer regarding the microphone type in real time.

Results of the model on real experimental data are presented in [12].

TABLE III
RESULTS OF INDEPENDENT TESTS

Test	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Class.	1	3	0	2	1	2	3	0	2	3	1	2	1	0	3	0
Accuracy	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

V. CONCLUSION

In this paper a complete explanation of the necessity for simulation data in the processing real photoacoustic measurement data is given. Software for simulations is designed based on the presented theoretical-mathematical model, while the credibility to the experiment is obtained using expert knowledge. Classification model for microphone type recognition is trained on the obtained database. Because of the huge, reliable dataset, knowledge about the photoacoustic experiment is embedded in the classification model so it could be optimized to a pretty simple topology, while the learning process was extremely efficient. In terms of precision and real time processing, classification model satisfies requirement of the photoacoustic experiment. In terms of reliability, classification model did not make any mistake in tests maintained with simulated data. The benefits of the presented model for PA measurements are multiple. By recognizing the microphone type the shape of transfer function and levels of signal exaggeration or attenuation are determined and that will simplify the further procedure of recognition of microphone characteristics in order to deprive PA signal of instrumental deviations. If the recognized microphone has flat characteristics the correction procedure is skipped. In the case of shaped response the correction procedure is done using only database of recognized microphone instead of whole database for all types of microphone. The processing time is saved this way. The generality of the model could be accomplished by extending the number of microphone types if such requirement of the experiment exists. In the future, we intend to explore modeling of noise distribution to generate data similar to real data and skip the first step of correction procedure, the noise removal.

REFERENCES

[1] Y. S. Abu-Mostafa, M. Magdon-Ismail, and H.-T. Lin, *Learning From Data*. AML Book, 2012.

[2] D. Allman, A. Reiter, and M. A. L. Bell, "Photoacoustic Source Detection and Reflection Artifact Removal Enabled by Deep Learning," *IEEE Trans. Med. Imaging*, vol. 37, no. 6, pp. 1464–1477, 2018.

[3] M. Lukić, Ćojbašić, M. D. Rabasović, D. D. Markushev, and D. M. Todorović, "Laser Fluence Recognition Using Computationally Intelligent Pulsed Photoacoustics Within the Trace Gases Analysis," *Int. J. Thermophys.*, vol. 38, no. 11, 2017.

[4] M. Lukić, Ž. Ćojbašić, M. D. Rabasović, and D. D. Markushev, "Computationally intelligent pulsed photoacoustics," *Meas. Sci. Technol.*, vol. 25, no. 12, p. 125203, 2014.

[5] M. N. Popovic, D. Furundzic, and S. P. Galovic, "Photothermal Depth Profiling Of Optical Gradient Materials By Neural Network," *Publ. Astron. Obs. Belgrade*, vol. 89, no. May 2015, 2010.

[6] S. P. Djordjevic, K.Lj., Markushev, D.D., Ćojbašić, Ž. M., Galović, "Photoacoustic measurements of the thermal and elastic properties of n-type silicon using neural networks," *Silicon , Springer*.

[7] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[8] S. Aleksic, D. Markushev, D. Pantic, M. Rabasovic, D. Markushev, and D. Todorovic, "Electro-acustic influence of the measuring system on the photoacoustic signal amplitude and phase in frequency domain," *Facta Univ. - Ser. Physics, Chem. Technol.*, vol. 14, no. 1, pp. 9–20, 2016.

[9] A. Handa, V. Patraucean, V. Badrinarayanan, S. Stent, and R. Cipolla., "SceneNet: Understand- ing real world indoor scenes with synthetic data.," in *CVPR*, 2015.

[10] A. Shrivastava, T. Pfister, O. Tuzel, J. Susskind, W. Wang, and R. Webb, "Learning from Simulated and Unsupervised Images through Adversarial Training," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2242–2251.

[11] G. Cosne *et al.*, "Using Simulated Data to Generate Images of Climate Change," in *ML-IRL workshop at ICLR*, 2020, pp. 1–9.

[12] M. I. Jordovic-Pavlovic *et al.*, "Computationally intelligent description of a photoacoustic detector," *Opt. Quantum Electron.*, vol. 52, no. 5, pp. 1–14, 2020.

[13] M. I. Jordović-Pavlović, M. M. Stanković, M. N. Popović, Ž. M. Ćojbašić, S. P. Galović, and D. D. Markushev, "The application of artificial neural networks in solid-state photoacoustics for the recognition of microphone response effects in the frequency domain," *J. Comput. Electron.*, vol. 19, no. 3, pp. 1268–1280, 2020.

[14] A. Rosencwaig, "Photoacoustic Spectroscopy of Solids," *Opt Commun, 7(4)*, pp. 305–308, 1973.

[15] L. C. M. Perondi, L. F, Miranda, "Minimal-Volume Photoacoustic Cell Measurement of Thermal Diffusivity: Effect of the Thermoelastic Sample Bending.," *J. Appl. Phys.*, vol. 62, pp. 2955–2959, 1987.

[16] M. Popovic, M. Nesic, S. Ciric-Kostic, M. Zivanov, D.Markushev, M. Rabasovic, S. Galovic, "Helmholtz Resonances in Photoacoustic Experiment with Laser-Sintered Polyamide Including Thermal Memory of Samples," *Int. J. Thermophys.*, vol. 37, no. 12, pp. 1–9, 2016.

[17] A. Rosencwaig and A. Gerscho, "Photoacoustic Effect with Solids: A Theoretical Treatment," *Science (80-.).*, vol. 190, pp. 556–557, Nov. 1975.

[18] F. A. McDonald and G. C. Wetsel, "Generalized theory of the photoacoustic effect," *J. Appl. Phys.*, vol. 49, no. 4, pp. 2313–2322, Apr. 1978.

[19] L. Rousset, F. Lepoutre, and L. Bertrand, "Influence of thermoelastic bending on photoacoustic experiments related to measurements of thermal diffusivity of metals," *J. Appl. Phys.*, vol. 54, no. 5, pp. 2383–2391, 1983.

[20] P. M. Nikolic and D. M. Todorović, "An investigation of semiconducting materials using a photoacoustic method. u: Technical sciences book 40, Monographs, Belgrade: Serbian Academy of Sciences and Arts Department, vol. DCXLVIII." Serbian Academy of Sciences and Arts, 2001.

[21] D. D. Markushev, M. D. Rabasović, M. V Nesic, M. N. Popovic, and S. P. Galovic, "Influence of thermal memory on thermal piston model of photoacoustic response," *Int. J. Thermophys.*, vol. 33, no.

10–11, pp. 2210–2216, 2012.

[22] M. V Nesic, S. P. Galovic, Z. N. Soskic, M. N. Popovic, and D. M. Todorović, "Photothermal thermoelastic bending for media with thermal memory," *Int. J. Thermophys.*, vol. 33, no. 10–11, pp. 2203–2209, 2012.

[23] S. P. Galovic, Z. N. Soskic, M. N. Popovic, Z. Stojanovic, D. Cevizovic, and Z. Stojanovic, "Theory of photoacoustic effect in media with thermal memory," *J. Appl. Phys.*, vol. 116, no. 2, pp. 0–12, 2014.

[24] D. M. Todorović, M. D. Rabasovic, and D. D. Markushev, "Photoacoustic elastic bending in thin film—Substrate system," *J. Appl. Phys.*, vol. 114, no. 21, p. 213510, 2013.

[25] D. M. Todorović, M. D. Rabasovic, D. D. Markushev, and M. Sarajlić, "Photoacoustic elastic bending in thin film-substrate system: Experimental determination of the thin film parameters," *J. Appl. Phys.*, vol. 116, no. 5, 2014.

[26] M. Popovic, *Generalizovani fotoakustični odziv dvoslojnih struktura*. Beograd: Zadužbina Andrejević, 2018.

[27] M. N. Popovic, "Fotoakustički odziv transmisione fotoakustičke konfiguracije i analiza rezonantnih fenomena za dvoslojne uzorke sa toplotnom memorijom," Univerzitet u Novom Sadu, 2016.

[28] M. N. Popovic, D. D. Markushev, M. V. Nesic, M. I. Jordovic-Pavlovic, and S. P. Galovic, "Optically induced temperature variations in a two-layer volume absorber including thermal memory effects," *J. Appl. Phys.*, vol. 129, no. 1, 2021.

[29] M. Popovic, M. Nesic, M. Zivanov, D. Markushev, and S. Galovic, "Photoacoustic response of a transmission photoacoustic configuration for two-layer samples with thermal memory," *Opt. Quantum Electron.*, vol. 50, p. 330, 2018.

[30] M. D. Rabasovic, M. G. Nikolic, M. D. Dramicanin, M. Franko, and D. D. Markushev, "Low-cost, portable photoacoustic setup for solid samples," *Meas. Sci. Technol.*, vol. 20, no. 9, p. 95902, 2009.

[31] M. Nesic *et al.*, "Development and comparison of the techniques for solving the inverse problem in photoacoustic characterization of semiconductors," *Opt. Quantum Electron.*, vol. 53, no. 7, p. 381, 2021.

[32] Q. V Le, J. Ngiam, A. Coates, A. Lahiri, B. Prochnow, and A. Y. Ng, "On Optimization Methods for Deep Learning," in *Proceedings of the 28th International Conference on Machine Learning*, 2011, pp. 265–272.

[33] Y. Bengio, "Practical recommendations for gradient-based training of deep architectures," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7700 LECTU, pp. 437–478, 2012.

[34] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, 2010, vol. 9, pp. 249–256.

[35] D. P. Kingma and J. Ba, "Adam: {A} Method for Stochastic Optimization," *CoRR*, vol. abs/1412.6, 2014.

Genetic Algorithm for Bent Functions Generating

Milan Stojanović and Suzana Stojković

Abstract— **The importance of unique bent functions (most significantly in cryptography) creates a demand for their generation. Bent function generation is an interesting problem and, in this paper, we explore the idea of using invariant spectral operations in a Genetic algorithm for generating bent functions. Invariant spectral operations, when executed on bent function, resulting function is also bent. If multiple operations are performed consecutively, then there is a possibility that the newly generated bent function is not unique. A genetic algorithm is used to search the solution space in order to produce the most unique bent functions, for the least number of invariant spectral operations.**

Index Terms— **Bent functions, invariant spectral operations, genetic algorithm.**

I. Introduction

Bent functions are Boolean functions most distant from affine functions. They were introduced by O.S. Rothaus in 1976. [1], and they have characteristics that are interesting for cryptographic applications. There are many algorithms for the generation of the bent function, see for example [2-9] and references therein.

A very important characteristic of the bent functions is flat Walsh spectra. All Walsh spectral coefficients of n-variable bent functions have the same absolute value equal to $2^{n/2}$. Invariant spectral operations are operations that do not change the absolute values of spectral coefficients, i.e., they only permute or change the sign of spectral coefficients. It follows that new bent functions can be generated from any known bent function by applying invariant spectral operations. References [7-9] elaborate methods for bent functions generation by using invariant spectral operations. The main disadvantage of those methods is that the same bent function can be generated by applying different sequences of operations.

Genetic algorithm is inspired by natural selection, that belongs to the evolutionary algorithm group. This algorithm is used to optimize a solution for a corresponding problem. It can be used most effectively when the search space is vast, but the solution does not need to be perfect, only optimal to some degree.

Milan Stojanović is master's-degree student in Computer Science at Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia (e-mail: milances.14@gmail.com)

Suzana Stojković is with Dept. of Computer Science, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14 18000 Niš, Serbia (e-mail: suzana.stojkovic@elfak.ni.ac.rs).

This paper proposes the usage of a Genetic algorithm for the generation of bent functions. Bent functions belong to the vast space of Boolean functions. Therefore, the search for unique bent functions can be presented as executing a sequence of invariant spectral operations, and optimization is used in the sequence of operations, so that we will produce as many different bent functions as possible.

The paper is organized in the following way: Section II presents the ANF representation of bent function. Section III covers Invariant spectral operations. Oscar-Bent functions are presented in Section IV and the Genetic algorithm is defined in Section V. Section VI explains the problem definition and usage of the Genetic algorithm for the generation of bent functions. Section VII goes over the results, and Section VIII gives a conclusion.

II. ANF Representation of Bent Functions

A. Definition

An n-variable Boolean function $f(x_1, x_2, \dots, x_n)$ can be presented by the algebraic normal form (ANF), or the positive polarity Reed-Muller expansion as:

$$f(x_1, \dots, x_n) = \sum_{i=0}^{2^n-1} S(i) \prod_{k=0}^{n-1} x_k^{i_k}$$

where $S(i)$ is the Reed-Muller spectral coefficient and $i_0 i_1 \dots i_{n-1}$ is the binary representation of the index i.

Reed-Muller spectral coefficients of bent functions are equal to 0 for each input vector with the number of ones greater than $n/2$. The maximal number of variables in a product term is called the degree of f [8].

B. Disjoint quadratic function

The disjoint quadratic function contains $n/2$ disjoint quadratic terms, defined as:

$$f(x_1, \dots, x_n) = x_1 x_2 \oplus x_3 x_4 \oplus \dots \oplus x_{n-1} x_n$$

III. Invariant Spectral Operations

A. Definition

Invariant spectral operations do not change the absolute values of Walsh spectral coefficients, they only permute or change the sign of spectral coefficients. These changes preserve the flat spectrum.

Due to the simplicity of invariant spectral operations in the Reed-Muller domain, all operations are introduced in this domain. For consistency, all examples will be provided starting from the Disjoint quadratic function for $n = 6$.

B. Function complement

Function complement is defined as:

$$f_2 = \overline{f_1} = f_1 \oplus 1$$

For example, if

$$f_1(x_1, x_2, x_3, x_4, x_5, x_6) = x_1x_2 \oplus x_3x_4 \oplus x_5x_6$$

The resulting function will be:

$$f_2(x_1, x_2, x_3, x_4, x_5, x_6) = x_1x_2 \oplus x_3x_4 \oplus x_5x_6 \oplus 1$$

C. Variable complement

Variable complement replaces the input variable i by its complement $x_i' = x_i \oplus 1$.

If variable complement on variable x_4 is performed, the function f_1 is transformed to:

$$f_2(x_1, x_2, x_3, x_4, x_5, x_6) = f_1(x_1, x_2, x_3, \overline{x_4}, x_5, x_6)$$
$$= x_1x_2 \oplus x_3(x_4 \oplus 1) \oplus x_5x_6$$
$$= x_1x_2 \oplus x_3x_4 \oplus x_3 \oplus x_5x_6$$

D. Disjoint spectral translation

Disjoint spectral translation replaces the input variable i by $x_i' = x_i \oplus x_j$, where $i \neq j$.

In the given example, if x_3 is replaced by $x_3 \oplus x_6$, following function is generated:

$$f_2(x_1, x_2, x_3, x_4, x_5, x_6) = f_1(x_1, x_2, x_3 \oplus x_6, x_4, x_5, x_6)$$
$$= x_1x_2 \oplus (x_3 \oplus x_6)x_4 \oplus x_5x_6$$
$$= x_1x_2 \oplus x_3x_4 \oplus x_4x_6 \oplus x_5x_6$$

E. Spectral translation

In the general case, we can define spectral translation as adding linear member x_i to the function:

$$f_2 = f_1 \oplus x_i$$

If in our example x_2 is added, resulting function is:

$$f_2(x_1, x_2, x_3, x_4, x_5, x_6) = x_1x_2 \oplus x_3x_4 \oplus x_5x_6 \oplus x_2$$

F. Permutation of variables

Permutation of variables is defined as the interchange of two input variables $x_i \leftrightarrow x_j$, where $i \neq j$.

$$f_2(x_1, \dots, x_i, \dots, x_j, \dots, x_n) = f_1(x_1, \dots, x_j, \dots, x_i, \dots, x_n)$$

In the given example if we interchange input variables x_3 and x_6 the resulting function is:

$$f_2(x_1, x_2, x_3, x_4, x_5, x_6) = f_1(x_1, x_2, x_6, x_4, x_5, x_3)$$
$$f_2(x_1, x_2, x_3, x_4, x_5, x_6) = x_1x_2 \oplus x_6x_4 \oplus x_5x_3$$

G. Generalized spectral translation

The generalized spectral translation is defined for the function f which has n variables $(n = 2 * k, k \geq 3)$ and contains $n/2$ disjoint quadratic terms:

$$f(x_1, .., x_n) = \cdots x_{i_1}x_{j_1} \oplus x_{i_2}x_{j_2} \oplus \dots \oplus x_{i_{n/2}}x_{j_{n/2}}$$

Performing generalized spectral translation on function f adds a new term $x_{k_1}x_{k_2} \dots x_{k_{n/2}}$ where

$$k_1 \in \{i_1, j_1\},\ k_2 \in \{i_2, j_2\}, \dots, k_{n/2} \in \{i_{n/2}, j_{n/2}\}.$$

If the starting function is f_1 is and if $k_1 = 1,\ k_2 = 3$, and $k_3 = 6$, resulting function f_2 is:

$$f_2(x_1, x_2, x_3, x_4, x_5, x_6) = x_1x_2 \oplus x_3x_4 \oplus x_5x_6 \oplus x_1x_3x_6$$

IV. Oscar-Bent functions

The bent function which does not have linear and constant members can be called Oscar-Bent function (the name derives from Oscar Rothaus, who first defined bent functions). For the bent function defined in (1), we can derive the Oscar-Bent function shown in (2) by using invariant spectral operations.

$$f_1(x_1, \dots, x_6) = x_1x_2 \oplus x_3x_4 \oplus x_5x_6 \oplus x_1 \oplus 1 \quad (1)$$
$$f_2(x_1, \dots, x_6) = x_1x_2 \oplus x_3x_4 \oplus x_5x_6 \quad (2)$$

To transform a bent function to its Oscar-Bent function we need to remove linear and constant members, which is done by using two invariant spectral operations: function complement and spectral translation. By counting only Oscar-Bent functions, we can deduce that the number of unique bent functions found with this algorithm is calculated by multiplying the number of Oscar-Bent functions with 2^{n+1}. The multiplier is found by calculating all possible combinations using two invariant operations mentioned above.

V. Genetic Algorithm

Genetic algorithm is a subclass of Evolutionary algorithm (EA), which is a subclass of Evolutionary computation and belongs to set of general stochastic search algorithm [10].

Population in both Genetic algorithms and in nature represents the set of individuals who are trying to survive and pass on their genes to the next generations. An individual can be interpreted as a set of genes and abilities, and how fit they are to survive in the current population and habitat. If we observe an individual as a solution to a problem, as well as in nature, optimization (survival of the fittest) will transpire, in the end, the fittest individual will represent an optimized solution to the problem. Given a population of individuals within some environment that has limited resources, competition for those resources causes natural selection (survival of the fittest). This in turn causes a rise in the fitness of the population. Given a quality function to be maximized, we can randomly create a set of candidate solutions, i.e., elements of the function's domain. We then apply the quality function to these as an abstract fitness measure – the higher the better. Based on these fitness values some of the better candidates are chosen to seed the next generation. This is done by applying recombination and/or mutation to them. Recombination is an operator that is applied to two or more selected candidates (the so-called parents), producing one or more new candidates (the children). The mutation is applied to one candidate and results in one new candidate. Therefore, executing the operations of recombination and mutation on the parents leads to the creation of a set of new candidates (the offspring). These have their fitness evaluated and then compete – based on their fitness (and possibly age) – with the old ones for a place in the next generation. This process can be iterated until a candidate with sufficient quality (a solution) is found or a previously set computational limit is reached [11].

The genetic algorithm can be described by the pseudo-code in Fig. 1.

```
InitializePopulation();
EvaluatePopulation();
while i < MaxIteration and
BestFitness < MaxFitness do
        Fitness = FitnessCalculation();
        Selection();
        ParentSelection();
        Reproduction();
        i++;
        BestFitness = Max(Fitness);
end while
return BestFitness
```

Fig. 1. Pseudo code detailing the genetic algorithm

A. Parent selection

Parent selection represents a strategy of selecting good parents to get a better next generation. The strategy should consist of some random chance in selection, so diverse parents will be used, and we can diverge from the local maximum (which can be reached by using the same group of parents).

- There are different strategies, for this paper, we have used:
- Roulette selection – odds of selection are determined by individual fitness and a corresponding piece of the roulette wheel is given; a random number is generated to represent a ball spin.
- Rang selection – is like Roulette selection, but fitness is scaled to give more chances to weaker individuals.
- Tournament selection – from a randomly selected group of individuals the best individual is chosen based on fitness.

B. Recombination and mutation

Recombination and mutation are used to produce a new solution to find the best one which solves the problem. Both methods may and may not be performed (based on chance which is determined on startup).

There are several recombination methods, but the most common is a crossover with one crossover point which is randomly selected. Genes from the first parent are copied to the crossover point, after which genes from the second parent are copied.

Mutation, if performed, results in randomly changing individual genes. For each gene, independently, it is determined whether the mutation will be performed or not.

C. Adult selection

Adult selection defines how will the new generation join the existing group of adults. Since the "habitat" can only sustain an already defined number of individuals, adult selection is needed to determine who will survive. Several methods are implemented:

- Full generational replacement – as the name applies, the parental generation is replaced with the new generation.
- Generational mixing – both generations are mixed, and the best of mixed generations survives.
- Overproduction – this method is a mixture between full generational replacement and generational mixing, in which the new generation has twice as many individuals as the parental generation, and only a half of the best child individuals survive to form the new parental generation.

Elitism can also be used, elitism enables keeping the best solution for the next solution, regardless of chosen adult selection.

VI. Genetic Algorithm For Bent Functions Generating

The problem can be defined by finding the most bent function by using the least number of invariant spectral operations. Since invariant spectral operations are performed on a bent function, a starting bent function needs to be defined. In our case, we start from the disjoint quadratic function.

For each implementation of a genetic algorithm, it is crucial to define an individual (which represents a solution to a problem) and a fitness function (which represents how good the solution is).

A. Individual representation

An individual is represented as a sequence of invariant operations which are performed on the most recent bent function.

B. Fitness function

It is recommended that the fitness function should be defined so it would have a minimum (the worst solution) and the maximum (the best solution), even though the boundaries can be arbitrary, the custom is to choose boundaries as 0 and 1.

To determine how good is the solution, we need to go back to the problem definition which states that we should find the most unique bent functions for the least number of invariant spectral operations. From this, we can derive that the fitness function can be calculated as the number of unique Oscar-Bent functions divided by the number of used invariant spectral operations.

By searching for the unique Oscar-Bent functions, we can generate the most bent functions, given that from the one Oscar-Bent function we can derive 2^{n+1} bent functions.

VII. Experimental Results

The application was developed in C#, and tests were performed on the laptop with the following configuration:

- CPU: Intel® Core™ i5-8250U CPU @ 1.6GHz
- RAM: 16 GB
- OS: 64bit Windows 10

Multiple parameters can be changed, and which can influence results (both performance and result wise). Testing all permutations of the possible combination of parameters is not a trivial task, and it is time-consuming. Therefore, some parameters were hardcoded with values that we perceived as best with our experience and using educated guesses.

Parameters that were hardcoded for all tests:

- Adult selection – Generational mixing

- Parent selection – Tournament selection
    - Tournament size – 20% of the population
- Possibility of gene mutation – 10%
- Possibility of recombination – 90%

A. Test 1 – Different number of genes

In this test we have chosen the number of variables to equal 6, population size is set to 10, and the number of generations is limited to 100. In this test, we will change the number of genes and compare the number of unique Oscar-Bent functions in an average of 5 runs. Results are shown in Table I.

TABLE I.
RESULTS OF TEST 1

Number of genes	Number of generated OBF	Time (s)
100	97.8	0.066
1 000	948.4	0.454
10 000	9 364.4	8.31
100 000	92 058	82.724

Through a different number of genes, we have seen that with linear growth of the number of genes, the number of unique Oscar-Bent functions grows in a linear fashion, with the growth factor between 9.5 and 10. When we analyze the time needed, it grows exponentially which is expected since the solution space grows exponentially as well.

B. Test 2 – Different number of variables

As in the previous test, the population size is set to 100, the number of generations is limited to 100 and the number of genes to 10 000. Here we will fluctuate number of variables. Results are shown in Table II.

TABLE II.
RESULTS OF TEST 2

n	Number of generated OBF	*Time (s)*
8	9 652.6	16
10	9 786.8	35.316
12	9 834.4	111.346
14	9 881.2	413.6

While an increasing number of variables we can observe that the number of unique Oscar-Bent functions increase with low percentages. Factor of growth for the time needed increases with each step, but it does not increase exponentially.

C. Test 3 – Application limits

In this test, we emphasized the performance limits of the application, not to the numbers we have, therefore we have run this test only once. Here, we have kept the number of genes to 10 000 and the number of generations to 100, as in the last test. But we have changed population size to 100. The results are shown in Table III.

TABLE III.
RESULTS OF TEST 3

n	Number of generated OBF	Time (s)
8	9 673	169.18
10	9 768	405.08s
12	9 845	1318.69s
14	9 877	1801.56s
16	N/A	N/A

VIII. CONCLUSION

We have seen that the usage of Genetic algorithm can be used in the generation of new bent functions. The performance of this approach indicates that future work can give promising results.

Memory is the biggest obstacle when working with bent functions. This problem can be approached by tracking only Oscar-Bent functions, which is performed in this paper. Further, all functions are kept in memory, which is a problem when expecting many unique Oscar-Bent functions, which we have seen in test 3. Future work will address this problem.

REFERENCES

[1] O. S. Rothaus, “On “Bent” Functions”, Journal of Combinatorial Theory, Series A, Vol. 20. No. 3, pp. 300-305, 1976.
[2] H. Dobbertin, “Construction of bent functions and balanced Boolean functions with high nonlinearity”, LCNS, vol. 1008, pp. 61-74, Springer, Berlin, Germany, 1995.
[3] J. Climent, F. Garcia, V. Requena, “On the iterative construction of bent functions”, in Proc. of the 5th WSEAS Int. Conf. on Inf. Security and Privacy (ISP06), pp. 15–18 World Scientific and Engineering Academy and Society (WSEAS), Wisconsin, USA, 2006
[4] J. Climent, F. Garca, V. Requena, “On the construction of bent functions of n+2 variables from bent functions of n variables”, Advances in Mathematics of Communications, vol. 2, pp. 421–431, 2008.
[5] C. Carlet, “A larger class of cryptographic Boolean functions via a study of the Maiorana McFarland construction”, LNCS, vol. 2442. pp. 549-564, Springer, 2002.
[6] H. Dobbertin, G. Leander, A. Canteaut, C. Carlet, P. Felke, P. Gaborit, “Construction of bent functions via niho power functions”, Journal of Combinatorial Theory, vol. 113, no. 5, pp. 779–798, 2006.
[7] M. Stanković, C. Moraga, R. Stanković, “Some Invariant spectral Operations for Functions with Disjoint Products in the Polynomial Form”, in Proc. EUROCAST 2017, LNCS, Vol. 10672, pp. 262-269, Springer, 2017.
[8] S. Stojković, M. Stanković, C. Moraga, R. Stanković, “Generation of Binary Bent Functions by Walsh Invariant spectral Operations Performed in Reed-Muller Domain”, Proceedings of 13th International Workshop on Boolean Problems, Bremen, Germany, pp. 255-266, September 19-21. 2018.
[9] R. S. Stanković, M. Stanković, C. Moraga and J. T. Astola, "Construction of Ternary Bent Functions by FFT-like Permutation Algorithms", 2020 IEEE 50th International Symposium on Multiple-Valued Logic (ISMVL), pp. 88-93, 2020.
[10] P. A. Vikhar, "Evolutionary algorithms: A critical review and its future prospects," 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC), Jalgaon, India, pp. 261-265, 22-24 Dec. 2016.
[11] A.E. Eiben, J.E. Smith, “Introduction to Evolutionary Computing”, 2nd ed., Berlin, Germany: Springer, 2015.

Application of Machine Learning Algorithms for Calculating Air Quality Index

Nebojša Bogdanović, Mladen Koprivica, *Member, IEEE*, Goran Marković, *Member, IEEE*

Abstract—**Air pollution is an ever-growing issue, especially severe in urban and industrial areas. Air Quality Index (AQI) is a unit of measuring the level of air pollution, which takes into account the concentrations of all relevant air pollutants. There are two main problems that must be addressed in AQI calculations, i.e. regression and classification. The regression problem consists of calculating (approximating) the AQI index based on the concentrations of different air pollutants. In classification problem, the measurements of air pollutants' concentrations are classified into different Air Quality Classes. In this paper a number of Machine Learning (ML) and Deep Learning (DL) algorithms were designed and used in order to solve both the regression and classification problems for AQI. The main goal was to present performance comparison for wide set of ML and DL algorithms based on the values of Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Coefficient of Determination (R squared) in regression tasks, and Accuracy in classification tasks. Also, the percentage of algorithms' convergence and the time needed to perform these regression and classification tasks are also measured.**

Index Terms—**Air Quality Index (AQI); Machine Learning; Deep Learning; Regression; Classification**

I. Introduction

Air pollution presents growing issue, which is especially severe in urban and industrial areas, and occurs whenever excessive quantities of pollutants such as gases, particulates, and bio-molecules are introduced into the atmosphere. It has harmful consequences on human population and other living organisms (i.e. it can cause diseases and/or even death, and impairing crops). Air pollutants can be solid particles, liquid droplets, or gases, and are classified as primary (i.e. directly emitted from the source) or secondary (i.e. formed in the atmosphere) pollutants. National environmental agencies set the standards and air quality guidelines regarding acceptable levels for air pollutants, while the air quality index (AQI) is used as an indicator in order to report the measuring of the air pollution and how unhealthy the air is (i.e. reports on possible associated health effects, above all for risk groups). AQI is calculated based on the maximum individual AQI measured for the observed criteria (air) pollutants, and this calculation is rather complex and thus its implementation is not suitable for applications with low-cost sensor platforms employed in the form of dense IoT-based sensor network. In fact the process of calculating AQI by formulas consists of two steps: (1) calculation of air quality index for every pollutant in each of the measurements separately, and (2) observing values of the all indexes for every measurement in order to find the maximum. On the other hand, in a case of application of machine learning (ML), a whole process of training and testing the algorithms takes longer than the use of formulas, but these algorithms only need to be trained once, before its application in real-time systems. Thus, when compared to formulas which needs to be used every time we have a different measurement, the time needed to perform this task by ML algorithms is shorter. It should be noticed, that formulas used for these calculations are not complex, but since the implementation of these formulas requires using multiple loops and case functions, the process takes longer when compared to the testing part of the ML and DL algorithms.

On the other hand, the more useful, flexible and scalable usage of AQI in terms of influence on the human population health, would be to deploy air quality forecasting system based on the measured levels of concentration of individual air pollutants, which would be able to predict AQI (i.e. air quality) locally and in short-term manner (hourly). This demands the use of dense network of low-cost sensors and thus requires simple solution for the determination of AQI based on the local low-quality air pollutant measurements.

So far, research community and environmental agencies have developed different methods for calculation of AQI, [1][2], but still no universally accepted method exist that is appropriate in all scenarios, [3]. The machine learning (ML) based methods are proposed as an obvious and natural solution for AQI determination and prediction, such as fuzzy lattices decision support system, [1], the support vector regression (SVR), [2], or different ML algorithms (linear regression, random forest, decision tree, SVR, and K-Nearest neighbor).

In this paper, the broad set of ML algorithms, including deep learning (DL), are observed as possible solutions for determination of AQI based on the measured levels of six criteria pollutants. Also, we here addressed two main issues in AQI calculation: regression problem that represents AQI calculation based on criteria pollutants concentrations, and classification problem in which the measurements of air pollutants' concentrations are classified into the Air Quality Classes. The output of the ML models is the approximation of the current values of AQI, while the prediction of the future values of AQI is something we are considering for the future works. In total, 8 different ML algorithms and 5 DL models were analyzed for the regression task, while 9 different ML algorithms and 3 DL models were observed for the classification task. We here observed much broader set of ML algorithms than in previous work, i.e. in [3]. ML algorithms and DL models were designed, optimized and tested based on dataset consisting of real-time measurements

Nebojša Bogdanović is a student at the School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11020 Belgrade, Serbia (e-mail: bn150118d@student.etf.bg.ac.rs).

Mladen Koprivica is with the University of Belgrade, School of Electrical Engineering, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mail: kopra@etf.bg.ac.rs).

Goran Marković is with the University of Belgrade, School of Electrical Engineering, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mail: gmarkovic@etf.bg.ac.rs).

gathered from 5 countries. The performance metrics are defined, and performance analysis and comparison of the observed ML algorithms and DL models is performed for regression and classification tasks.

The paper is organized as follows. In the section II the basic concepts related to air quality pollutants, monitoring, and scale are given. Section III gives short description of the observed machine learning algorithms and the deep learning models observed in this paper, as well as a short description of AQI regression and classification tasks, while a dataset used in ML and DL algorithm training and performance analysis is described in section IV. The main results and conclusion are presented in section IV, followed by the final concluding remarks.

II. Air Pollutants and Air Quality Scale

Air pollution is most frequently man-made. It usually comes from factories, powerplants and heating plants which use unrenewable energy sources, cars and public transport.

According to International Energy Agency (IEA) [4], from the year 2018 to 2040 the projected energy demand should rise annually by 1.3%. This projected growth can be seen on Fig. 1, where in the year 2020 around 60% of all the energy should be generated using non-renewable energy sources. While the use of renewable energy sources is projected to increase by the year 2040, because of the growth in energy demand, the amount of energy generated by coal, gas, oil and nuclear energy will not decrease.

This is a growing problem, in both the developed and in countries in development, because a usage of fossil fuels results in high concentrations of air pollutants released into the atmosphere. Many of developed countries are fighting this problem by imposing laws which are restricting the amounts of fossil fuels burned each year. Also, powerplants and factories are required to use filters in order to reduce the emission of air pollution into the atmosphere.

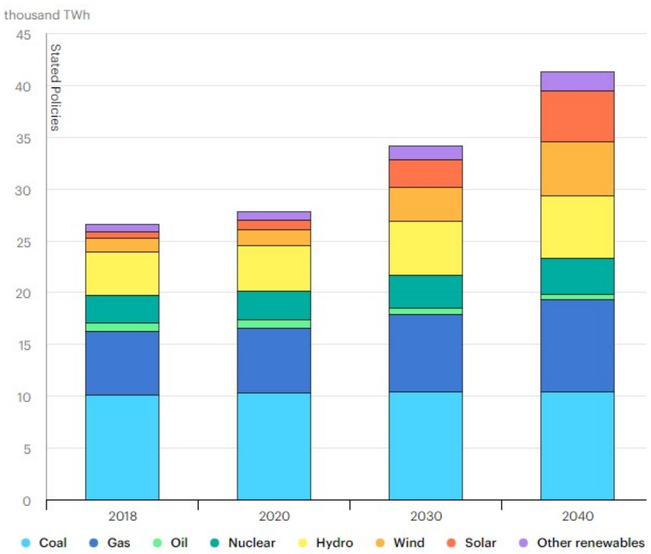


Fig. 1. Projected growth of energy demand from 2018 to 2040, [4]

The most common types of air pollution according to New South Wales Ministry of Health (NSW Health) [5], are listed below:

- Carbon Monoxide (CO), mostly generated by motor vehicles and industry plants;
- Ozone (O_3), the main component of smog, a product of interaction between sunlight and emissions from motor vehicles and industry plants;
- Particulate Matter (PM 2.5, PM 10), the small solid particles and liquid droplets suspended in air, made up of variety of components including nitrates, sulfates, organic chemicals, metals, soil or dust particles and allergens, which mostly comes from motor vehicles and industry plants;
- Nitrogen Dioxide (NO_2), generated by motor vehicles, industry plants, and unflued gas-heaters; and
- Sulphur Dioxide (SO_2), generated by the fossil fuel combustion at power plants and industrial facilities.

A. *Air Quality Scale*

Air Quality Index (AQI) can be calculated in a number of different ways, and depending on which formulas are used for calculations, there are different AQI scales. In this paper, the formulas and scales used are created by the Central Pollution Control Board, Ministry of Environment, Forests and Climate Change in India. This corresponding air quality scale is presented in Table I.

TABLE I
Air Quality Scale (India)

Category	AQI Index	Possible Health Impacts
Good	0-50	Minimal health impacts
Satisfactory	51-100	Minor breathing discomfort to sensitive people
Moderate	101-200	Breathing discomfort to the people with lung, asthma and heart diseases
Poor	201-300	Breathing discomfort to most people on prolonged exposure
Very Poor	301-400	Respiratory illness on prolonged exposure
Severe	401-	Affects healthy people and seriously impacts those with existing diseases

III. Algorithms and Problems

As defined in the introduction, there are two types of challenges in calculating AQI index which are addressed in this paper, the regression and the classification tasks. Both of these issues hold valuable information when calculating levels of air pollution. Some of the reasons for using ML algorithms in this area are:

- Provision of real-time decision support for air quality sensors, especially in a case of wide usage of low-cost sensors (i.e. for IoT-based environmental monitoring networks). Specifically, a verification that sensors for monitoring concentrations of various pollutants are working well, to predict the missing values in a case of sensor malfunction, and to evaluate inputs and decide whether and alarm should be triggered or not [1];
- Improvement of sensor performance for lower-cost air quality monitoring [6]; and
- Forecasting (prediction) of future values of pollution concentrations and AQI index [2] [3].

In this paper, we have performed comparison of the wide set of various machine learning algorithms for regression and classification tasks, such as: Multiple Linear Regression (MLR), Stochastic Gradient Descent (SGD) Classifier based on Linear Regression, Support Vector Machine (SVM), K-

Nearest Neighbors (KNN), Random Forest (RF), Decision Tree, Extra Trees Regression, Adaptive Boosting based on Decision Trees (AdaBoost), and Gradient Tree Boosting (GradBoost). Also, we designed and estimated performance for several Deep Learning algorithms in both tasks

Regression and classification tasks are rather similar, and thus the classification task can be realized by classifying the results achieved by regression algorithms into the respective categories in Table I. In this case, we would achieve 100% accuracy for classification task for the all algorithms, except for Multiple Linear Regression, but a time needed to execute would be even higher than for the regression algorithms. This is why in this paper we proposed a different method. In our method, the input data used for classification algorithms is the same as for the regression algorithms, and that is just the concentrations of the pollutants of measurements, while the labels used for training of the ML algorithms and DL models are final categories in Table I. By using this method we expected slightly lower classification accuracy (which will be discussed in section VI), when compared to the first method (based on regression), but the time needed to execute such algorithms would be, depending on an algorithm, from two to ten times smaller than for their respective regression algorithms.

The performance metrics for AQI regression algorithms were the values of Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Coefficient of Determination (R^2), while for the classification tasks we used the algorithm accuracy as the main performance metric. As the additional performance metrics for both tasks, we measured percentage of convergence and the elapsed time needed to perform these tasks for all observed algorithms.

IV. Air Pollutant Measurement Dataset

The dataset used in the analysis was created from data gathered from websites data.world [7], and openaq.org [8], and it consists of 35440 independent measurements from 5 countries (Serbia, India, USA, Australia and Taiwan). The measurements data were gathered from 2016 to April 2021. Each measurement consists of measured concentrations of Carbon-monoxide (CO), Ozone (O_3), particulate matter PM 2.5 and PM 10, Nitrogen-dioxide (NO_2) and Sulphur-dioxide (SO_2). The concentrations of all of the pollutants are measured in μg/m^3, except for CO, which is measured in mg/m^3. The mean values (mean) and standard deviations (std) of the measurements in dataset, are given in Table II.

TABLE II
Dataset description

	Count	Mean	Std
CO	35440	1.39088	1.34341
O_3	35440	51.9751	61.5998
PM 2.5	35440	76.7299	112.159
PM 10	35440	152.993	185.914
NO_2	35440	51.9291	56.3290
SO_2	35440	8.51487	8.71780
AQI	35440	202.835	195.820

Based on these concentrations, the reference AQI index and the air quality class were calculated for each of the measurements, by using formulas implemented in Python scripts. The dataset was divided into training and test sets in the ratio 90%:10%, and the training set was further divided into training and validation sets in the same ratio.

V. Results of Performance Analysis

The performance analysis of observed machine learning algorithms is performed by using Scikit-learn library for the Python programming language, while the deep learning algorithms were implemented using Tensorflow and Keras libraries for Python. The implementations were executed on Google Colaboratory cloud computing platform, by using Intel(R) Xeon(R) CPU @ 2.30 GHz processing unit with 16 GB of available RAM memory.

Both the machine learning and deep learning algorithms were trained and tested independently 42 times, with the values for *random_state* parameter ranging from 0 to 41, in order to guarantee different train/test splits of the dataset for the each iteration of the observed algorithm.

A. Regression algorithms

The analysis showed that all of the regression algorithms have the convergence rate of 90.47% (38/42). In 4 executions where the algorithms diverge, the corresponding MAE and RMSE values were not taken into account in the calculation of the mean values and standard deviations of these errors.

The five Deep Learning models, marked DL#1 to DL#5, were designed, optimized and used. These neural networks models are defined as: DL#1 model with 3 hidden layers comprising of with 128 neurons in each layer, DL#2 model with 3 hidden layers with 256 neurons in each layer, DL#3 model with 3 hidden layers with 512 neurons in each layer, DL#4 model with 3 hidden layers with 128 neurons in the first hidden layer, 1024 neurons in the second hidden layer, and 128 neurons in the third hidden layer (DL#4), and DL#5 model with 3 hidden layers with 256 neurons in the first hidden layer, 1024 neurons in the second hidden layer, and 256 neurons in the third hidden layer. Data is normalized in input layer of each DL model. The activation function for all layers was a ReLU function, while the loss function used was Mean Squared Error (MSE). The Adam optimization function was used with the learning rate of 0.001, and every neural network model is trained over 100 epochs. Different numbers of epochs for training DL models were considered during the design of these models, for both regression and classification tasks. In this process, it is observed that even if for some of the lower numbers of epochs the algorithms performed similarly as for 100 epochs, the results were not consistent enough, e.g. the convergence rate was lower (i.e. for 80 epochs the algorithms converged in 28/42 cases). Thus, we choose the number of 100 epochs, since the further rise in the number of epochs did not give better results.

The mean values and standard deviations of MAE and RMSE for all of the algorithms used in AQI regression task are shown in Table III, while in Table IV the times needed to train (single execution) of all these algorithms are given. The time needed for execution of all trained algorithms for one test measurement was similar and very short (in ms).

From the MAE and RMSE values, shown in Table III, it can be inferred that the best overall performance in a case of regression was achieved by using Adaptive Boosting based on Decision Trees. Also, by analyzing tables III and IV it can be inferred that the simpler algorithms, such as Multiple Linear Regression and Decision Tree take shortest time to train (and execute). On the other hand, the algorithms that consist of a large number of decision trees (i.e. Random Forest, AdaBoost or Extra Trees), SVM and deep learning algorithms, take the longest time to train (and execute) due to the complexity.

TABLE III
REGRESSION ALGORITHMS - MAE AND RMSE VALUES

Algorithm	MAE		RMSE	
	Mean	Std	Mean	Std
Random Forest	0.515066	0.065481	4.202714	1.167030
Decision Tree	0.652400	0.094673	5.911044	1.294427
AdaBoost	0.102419	0.02371	1.056525	0.477244
GradBoost	0.492849	0.049863	3.154737	0.952986
Extra Trees	0.469917	0.043560	2.800131	0.763884
KNN	7.586408	0.243762	17.655297	1.181481
SVM	6.811546	0.204737	12.891746	1.510237
MLR	35.95306	0.51608	52.938264	1.547197
DL#1	1.857936	0.407002	3.732621	0.413279
DL#2	1.552141	0.340354	3.413873	0.465218
DL#3	1.819313	0.501331	3.687509	0.576523
DL#4	1.616337	0.356425	3.417748	0.458047
DL#5	1.719751	0.539735	3.565049	0.655192

TABLE IV
REGRESSION ALGORITHMS - DURATION OF TRAINING (SINGLE EXECUTION)

Algorithm	Time [s]		Alg.	Time [s]	
	mean	std		mean	std
Random Forest	142.762	8.4116	MLR	0.007	0.0112
Decision Tree	0.215	0.0043	DL#1	170.780	27.4982
AdaBoost	69.558	2.3812	DL#2	250.169	16.4083
GradBoost	42.799	1.1223	DL#3	494.753	20.3062
Extra Trees	103.272	2.2578	DL#4	354.726	24.7223
KNN	0.499	0.0156	DL#5	627.936	30.0591

Furthermore, a more detailed statistical and error analysis (i.e. the minimum and maximum error values, the threshold values corresponding to 25%, 50% and 75% of instances), as well as time needed for training of AdaBoost algorithm are shown in Table V.

TABLE V
DETAILED ANALYSIS OF ADABOOST ALGORITHM FOR REGRESSION TASK

	MAE	MSE	RMSE	R^2	Time [s]
Mean	0.10242	1.338014	1.056525	0.999965	69.558
Std	0.02371	1.122726	0.477244	0.000029	2.3812
Min	0.06659	0.178894	0.422959	0.999895	64.468
25%	0.08444	0.40364	0.635299	0.999943	67.793
50%	0.09975	0.807562	0.897779	0.999978	69.868
75%	0.11428	2.176002	1.474417	0.999989	71.291
Max	0.15632	3.988713	1.997176	0.999995	75.529

As obvious in Table V, 50% of the MAE values for AdaBoost algorithm are under 0.1, with its mean value being just over 0.1. These are by far the best values of MAE for all of the observed regression algorithms that were compared in this paper.

B. *Classification algorithms*

The analysis showed that all the observed classification algorithms have the convergence rate of 100%, which means that the algorithms manage to converge around the mean values of accuracy in all of the 42 independent executions. Besides machine learning algorithms, three Deep Learning models, marked DL#6 to DL#8, were designed, optimized and used. These neural networks models are defined as: DL#6 model with 2 hidden layers with 128 neurons in each layer, DL#7 model with 2 hidden layers with 256 neurons in each layer, and DL#8 model with 2 hidden layers with 512 neurons in each layer. Data is normalized in the input layer of every neural network model. The activation function of each hidden layer is the ReLU function and the activation function of the output layer is the Softmax function. The loss function is Binary Cross-entropy, and the metrics of the loss function is the binary accuracy function with the 0.5 threshold value. The Adam optimization function was used with the learning rate of 0.001, and every neural network model is trained over 100 epochs.

The mean values and corresponding standard deviations (std) of classification accuracy for all observed classification algorithms, as well as the times needed for the training are given in Table VI.

Based on accuracy values for different algorithms, shown in Table VI, it can be inferred that the best algorithm for the classification task is the Gradient Tree Boosting. Also, the difference in time needed to train (and execute) more and less complex classification ML algorithms is not as big as it is in case of regression algorithms. This can be explained by the fact that the classification problem is easier to solve, and it does not require as much time as the regression one. Deep learning algorithms for classification take longer to train, since these were trained over 100 epochs.

TABLE VI
CLASSIFICATION ALGORITHMS - ACCURACY AND DURATION OF TRAINING

Algorithm	Accuracy		Time [s]	
	mean	std	mean	std
Random Forest	0.998683	0.00052	26.04913	0.303142
Random Forest Hybrid	0.998388	0.000612	18.30846	0.115094
Decision Tree	0.99822	0.000775	0.096139	0.003632
AdaBoost	0.998233	0.000753	0.104626	0.003506
GradBoost	0.999422	0.000432	24.62098	3.891909
Extra Trees	0.990682	0.001651	12.96546	0.247514
KNN	0.92313	0.004718	0.232663	0.008677
SVM	0.976849	0.00237	67.03444	6.434671
SGD	0.695058	0.009976	0.444438	0.019513
DL#6	0.994421	0.001118	118.5808	22.9412
DL#7	0.994591	0.001186	139.3796	2.512952
DL#8	0.994536	0.001218	440.6514	24.03057

The more detailed analysis of the Gradient Tree Boosting algorithm is shown in Table VII (the minimum and the maximum accuracy values are given, as well as threshold values corresponding to 25%, 50% and 75% of instances, and time needed for training), while estimated confusion matrix for this algorithm is shown on Fig. 2. It can be seen that in a case of Gradient Tree Boosting algorithm, only 3 of 7088 independent measurements of the test set used were misclassified (see confusion matrix in Fig. 2).

When compared to the classification results achieved in [3], we here achieved slightly better results for classification accuracy for the same algorithms that were used in both papers. However, in this paper the number of epochs for training the DL algorithms was higher than in [3], which can be one of the reasons for better accuracy results. Yet, the novelty of our paper, when compared to the work in [3], is that we included a number of algorithms that were not implemented in [3], for which we here achieved even better results in classification accuracy.

TABLE VII
DETAILED ANALYSIS OF GRADBOOST ALGORITHM

	Accuracy	Time [s]
mean	0.999422	24.62098
std	0.000432	3.891909
min	0.998025	22.1305
25%	0.999154	22.70442
50%	0.999436	22.99523
75%	0.999718	23.20253
max	1	35.08803

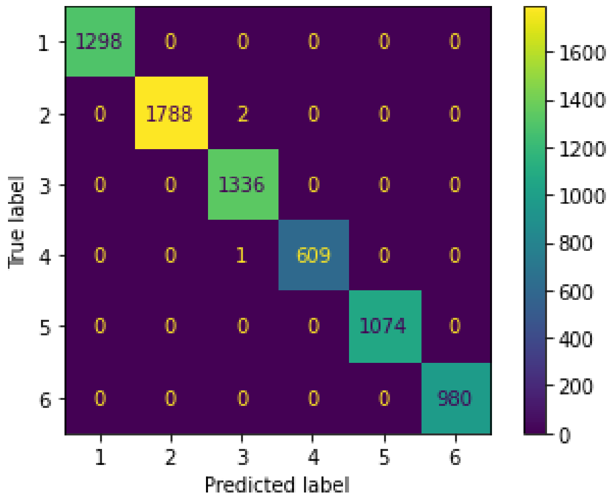


Fig. 2. Confusion matrix for the GradBoost algorithm

VI. CONCLUSION

Alongside global warming, the air pollution is one of the most alarming global ecological problems. Thus, developed countries, international health organizations, as well as some international companies are investing money in to reduce the impact air pollution have on global health. Also, some of air pollution aware companies try to motivate people to contribute to the cause, by giving them a chance to connect their air quality sensors with the global network of sensors, created by these companies. I.e., one of the most famous companies and websites that does this is called IQ Air [8].

The main topics covered in this paper are calculating the AQI (regression task), and the classification of air pollutant measurements into different air quality classes. We observed a wide set of machine learning and deep learning regression and classification algorithms for these tasks, and presented the performance comparison of these algorithms, based on the values of MAE, RMSE and accuracy, as well as the time needed to execute these algorithms. In total, 8 and 9 ML algorithms, as well as 5 and 3 DL models, were observed for regression and classification tasks, respectfully. It is shown that the AdaBoost algorithm presents best choice in the case of regression task, while the GradBoost algorithm presents the best choice in the case of classification tasks.

The presented results, as well as the designed and trained algorithms, present a foundation of a forecasting model, for predicting the missing and future pollution measurements and values of air quality index. This forecasting model could be used as a part of mobile application, which would inform users about the daily and weekly predictions of the pollution levels. This is one of the ideas for the future works. Another possible way of using the designed and trained algorithms, would be implementing in industrial plants.

ACKNOWLEDGMENTS

This work has been partly supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

REFERENCES

[1] I. N. Athanasiadis, V. G. Kaburlasos, P. A. Mitkas, and V. Petridis, “Applying Machine Learning Techniques on Air Quality Data for Real Time Decision Support”, 1st Intl. Symposium on Information Technologies in Environmental Engineering (ITEE 2003), pp. 51, 2003, ICSC/NAISO Academic Press.

[2] M. Castelli, F. M. Clemente, A. Popovič, S. Silva, and L Vanneschi, "A Machine Learning Approach to Predict Air Quality in California", Complexity, vol. 2020, Article ID 8049504, 23 pages, 2020.

[3] M. Sharma, J. Samyak, S. Mittal, and T. Sheakh, “Forecasting and Prediction of Air Pollutants Concentrates Using Machine Learning Techniques: The Case of India”, IOP Conference Series: Materials Science and Engineering, ICCRDA 2020, Vol. 1022, 012123, 2021.

[4] Electricity generation by fuel and scenario, 2018-2040, IEA, Paris https://www.iea.org/dataand-statistics/charts/electricity-generation-by-fuel-and-scenario-2018-2040 (last time accessed on 10.06.2021.)

[5] https://www.health.nsw.gov.au/environment/air/Pages/common-air-pollutants.aspx (last time accessed on 10.06.2021.)

[6] N. Zimmerman, A. A. Presto, S. P. N. Kumar, J. Gu, A. Hauryliuk, E. S. Robinson, A. L. Robinson and R. Subramanian, "A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring", Atmos. Meas. Tech., vol. 11, no. 1, 2018, pp. 291–313.

[7] https://data.world/ (last time accessed on 10.06.2021.)

[8] https://openaq.org/ (last time accessed on 10.06.2021.)

СПЕЦИЈАЛНА СЕСИЈА - СТОХАСТИЧКЕ МЕТОДЕ У МЕРЕЊИМА
/
SPECIAL SESSION - STOCHASTIC METHODS IN MEASUREMENT

(СС-МЛ/SS-MLI)

Merenje snage i energije vetra anemometrom bez pokretnih delova

Boris Ličina, Bojan Vujičić, Platon Sovilj, Vladimir Vujičić

***Apstrakt* — U radu se analizira merenje snage i energije vetra anemometrom sa nepokretnim ramovima. Razmatraju se dva scenarija – merenje standardnom sampling metodom (SSM) i merenje dvobitnom stohastičkom digitalnom mernom metodom (SDMM). Novina je softversko diterovanje izlaza iz anemometra.**

***Ključne reči* — snaga vetra; energija vetra; dvobitna SDMM; anemometer sa nepokretnim ramovima.**

I. Uvod

U radu [1] je pokazano kako se anemometrom sa šoljicama može vrlo tačno meriti energija vetra. Ta činjenica je posebno bitna u istraživanju izdašnosti lokacija za vetro-parkove. Merenje prosečnog smera vetra je daleko jednostavniji problem – prosto se standardni vetrokaz mehanički spregne sa apsolutnim enkoderom [2] u cilju merenja ugla i direktno se dobija digitalna informacija o trenutnom smeru vetra, a usrednjavanjem tih informacija se dobija glavna – prosečan smer vetra. Otežavajuća okolnost je primena oba senzora, i anemometra sa šoljicama i vetrokaza, što su im najvažniji delovi pokretni pa u teškim vremenskim uslovima drastično opada njihova upotrebljivost.

Kao jedno od mogućih rešenja da se navedena otežavajuća okolnost prevaziđe je projektovan i realizovan senzor za istovremeno merenje brzine i smera vetra – anemometer sa nepokretnim ramovima [3]. Tema ovog rada je primena anemometra sa nepokretnim ramovima u merenju energije i prosečnog smera vetra na potencijalnim lokacijama za izgradnju vetroparkova.

II. Postavka problema

Konstatujmo, najpre, da je u [1] opisano vrlo tačno merenje energije vetra primenom dvobitne SDMM. Konstatujmo, dalje, da je u [4] pokazano kako se dvobitnom SDMM može tačno meriti i u prisustvu značajnih nelinearnosti primenjenog senzora.

Boris Ličina – Univerzitet Privredna Akademija Novi Sad, Fakultet za primenjeni menadžment, ekonomiju i finansije, Jevrejska 24/1, 11000 Beograd (e-mail: boris.licina@yahoo.com).

Bojan Vujičić – Univerzitet u Novom Sadu, Fakultet tehničkih nauka Novi Sad, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: bojanvuj@uns.ac.rs).

Platon Sovilj – Univerzitet u Novom Sadu, Fakultet tehničkih nauka Novi Sad, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: platon@uns.ac.rs).

Vladimir Vujičić – Univerzitet u Novom Sadu, Fakultet tehničkih nauka Novi Sad, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: vujicicv@uns.ac.rs).

Konstatujmo još da je izlaz iz anemometra sa nepokretnim ramovima digitalan (zapravo ima dva digitalna izlaza – jedan je brzina vetra, a drugi smer vetra), Sl. 1. i da su date dve kalibracione nelinearne krive [3], Sl. 2. i Sl. 3.

Problem koji se rešava je:

a) Izmeriti tačno, u dugom vremenskom intervalu, energiju vetra na datoj lokaciji uzimajući u obzir nelinearnost anemometra sa nepokretnim ramovima. Problem je ekvivalentan sa merenjem srednje snage vetra u istom vremenskom intervalu.
b) Izmeriti tačno srednji smer vetra u istom vremenskom intervalu kao pod a).

III. Predlog rešenja

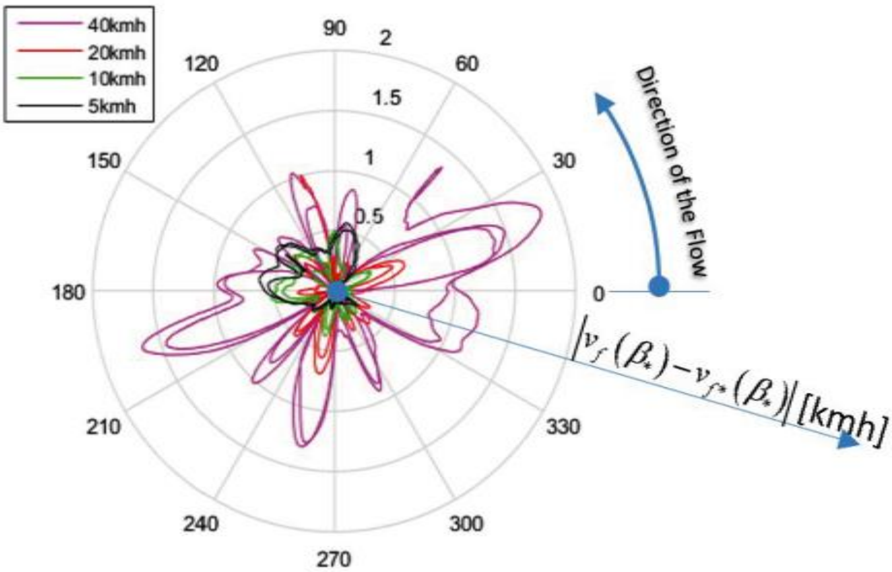
Činjenica da senzor daje dva digitalna izlaza prirodno nameće primenu dvobitne SDMM gde se hardversko analogno diterovanje zamenjuje softverskim diterovanjem [5], a dalja obrada je, principjelno, ista kao u [1] i u [4]. Šta više, s obzirom da je softversko diterovanje drastično jednostavnije od hardverskog i, uz to, i adaptibilno, moguće je primeniti i optimalnu rezoluciju od 3 bita [6] i time, uz minimalno komplikovanje obrade, dobiti efektivno ubrzanje od 9 puta postojeće tehnologije merenja. Pri tome se ne narušava tačnost merenja koja je definisana kalibracionim krivama. Rad [4] pokazuje kako se koristi dvobitna SDMM u merenju nelinearnim senzorom. Osnova je imati na raspolaganju ili:

i. inverzni kalibracioni polinom [4] ili
ii. dijagram greške odgovarajuće merne veličine [7]

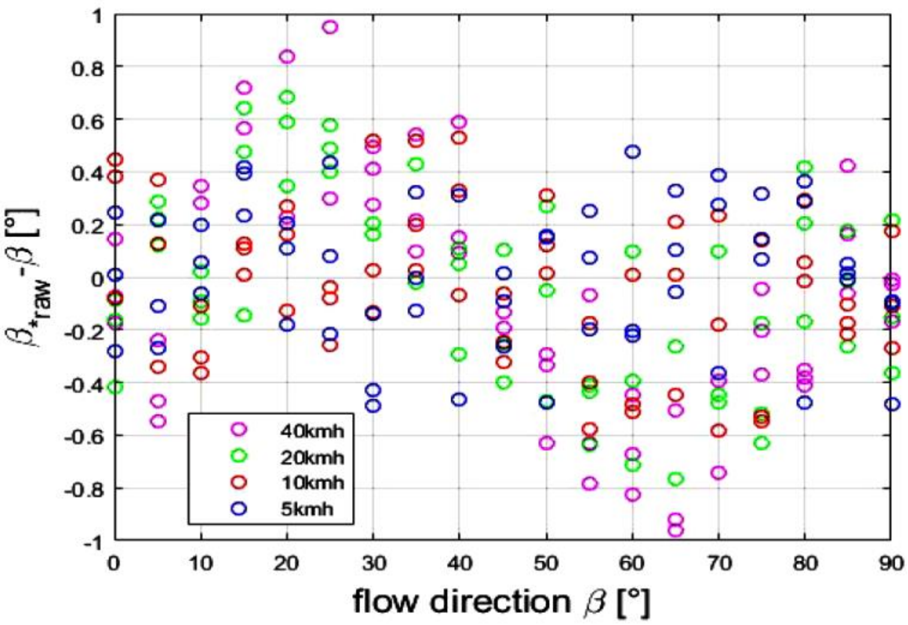
U [7] je pokazano kako se mogu meriti dve uzajamno spregnute veličine – prenosni odnos i fazna greška (ugao). U našem slučaju ovde, uzajamno spregnute veličine su: jačina (amplituda) vetra i smer (ugao) vetra. Novi princip nelinearne regresije, razvijen u [7], omogućuje korišćenje izlaza iz nelinearnog senzora, strujnog mernog transformatora (SMT) u [7], odnosno anemometra sa ramovima u ovom slučaju, do granice tačnosti kalibracionog sistema kojim su dobijene krive na Sl. 2. i Sl. 3. bez ikakve potrebe za linearizacijom senzora.

Sa druge strane, digitalni izlazi anemometra sa ramovima omogućuju i direktnu obradu primenom SSM i dobijanje traženih veličina – energije i srednjeg smera vetra. Problem koji se, generalno, tu javlja je eleboriran u [8] a koji se svodi na ogromno uvećanje greške usled integralne nelinearnosti primenjenog ADC u senzoru. Taj problem je, kako je pokazano, za bar dva reda veličine manji u slučaju primene dvobitne SDMM. Stoga je, bez sumnje, potrebno i ovde primeniti softverski diterovanu dvobitni SDMM, a ne direktno SSM.

Sl. 1. Fotografija evolucije anemometra sa ramovima [3].



Sl. 2. Kalibraciona kriva brzine vetra [3].

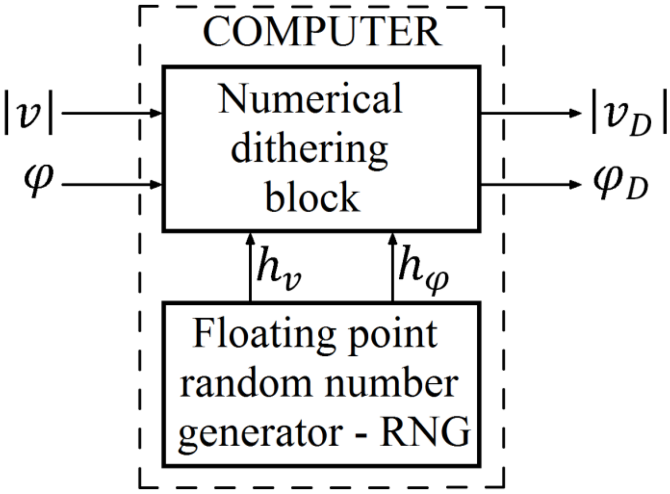


Sl. 3. Kalibraciona kriva smera vetra [3].

Senzor bez pokretnih delova korišćen u [3] na svom izlazu daje dva digitalna podatka u floating point reprezentaciji:

- intenzitet (brzinu) vetra $|v|$ i
- smer duvanja vetra φ

Kompletna obrada pomoću dvobitne SDMM obavlja se u mikroračunaru (Beaglebone, Raspberry), a njena opšta šema prikazana je na Sl. 4.

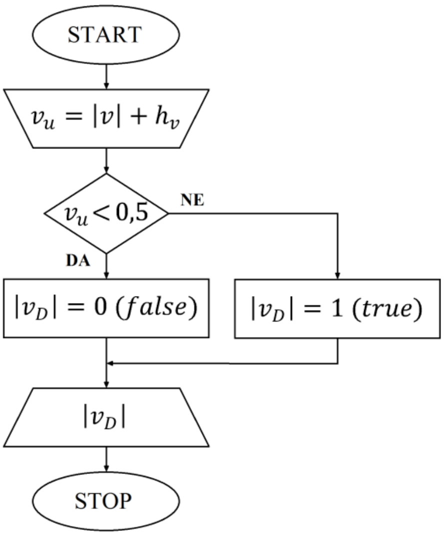


Sl. 4. Opšta šema numeričkog diterovanja

Sl. 5. prikazuje algoritam računanja jednobitnog diterovanog intenziteta (brzine) vetra $|v_D|$, gde se na ulazni signal brzine vetra $|v|$, dobijen sa senzora bez pokretnih delova, superponira diterski signal h_v, tj: $v_u = |v| + h_v$, pri čemu je:

$$|v| \leq V_{FS} \tag{1}$$

$$|h_v| \leq \frac{V_{FS}}{2} \tag{2}$$

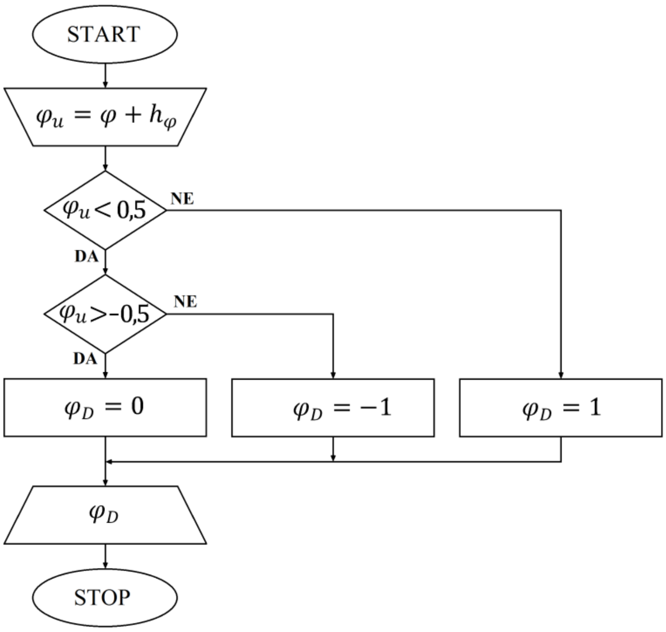


Sl. 5. Jednobitni diterovani intenzitet vetra $|v|$

Sl. 6. prikazuje algoritam računanja dvobitnog diterovanog smera duvanja (ugla) vetra φ_D, gde se na ulazni signal smera duvanja (ugla) vetra φ, takođe dobijenog sa senzora bez pokretnih delova, superponira diterski signal h_φ, tj: $\varphi_u = \varphi + h_\varphi$, pri čemu je:

$$|\varphi| \leq \pi \tag{3}$$

$$|h_\varphi| \leq \frac{\pi}{2} \tag{4}$$



Sl. 6. Dvobitni diterovani smer duvanja (ugao) vetra φ_D

Napomenimo još da je postupak softverskog diterovanja prikazan samo za mantise ulaznih podataka. Operacije nad eksponentima nisu prikazane jer su mnogo jednostavnije i, u osnovi, definisane relacijama (1) i (2), odnosno (3) i (4), respektivno.

IV. DISKUSIJA

U radu [8] je pokazano da nelinearnost dvobitnog stohastičkog adicionog AD konvertora (SAADK), ukoliko se koristi DAC AD5791, praktično, na dugom vremnskom intervalu (recimo jedan dan) pravi grešku reda ppm, a greška obrade praktično iščezava [4]. Ukoliko se, pak, koristi SSM i ADC (recimo 16-bini ADC LTC1605 koji je obrađen u [8]) čija je integralna nelinearnost reda 16 ppm, što je korektna ocena, onda je srednja vrednost kuba izlaza iz ADC-a:

$$(A + \delta A)^3 = A^3 + 3A^2(\delta A) + 3A(\delta A)^2 + (\delta A)^3 \tag{5}$$

gde je sa A označen moduo brzine, a sa δA integralna nelinearnost primenjenog AD konvertora.

U prvoj aproksimaciji srednja vrednost kuba izlaza iz ADC-a iznosi:

$$3 \cdot \frac{(\delta A)}{A} = 3 \cdot 16\,ppm = 48\,ppm \tag{6}$$

dakle, ona je za gotovo dva reda veličine veća od one dobijene pomoću dvobitnog SAADK koji ima ofset analognog sabirača od 1/5000 FS (1 mV, na 5 V). Ako kao ofset analognog sabirača prihvatimo vrednost δA, vrednost unutar LSB datog AD konvertora, kako je to pokazano u [4], tada je njegova vrednost još manja i iznosi $1/2^{16} = 1/64000$.

U zavisnosti od metode primenjene u konstrukciji ACD u SSM metodi, može se značajno smanjiti uticaj integralne nelinearnosti, ako se koristi samo MSB u dvobitnoj softverski diterovanoj SDMM u ovom slučaju. Ili MSB i prvi sledeći niži bit u slučaju trobitne softverski diterovane SDMM što je, dokazano je u [6], optimalna rezolucija u primeni SDMM. Naime, u slučaju da je, na primer, primenjena metoda sukcesivnih aproksimacija u konkretnom ADC-u na integralnu nelinearnost ne utiču svi otpornici u R-2R mreži, nego samo prva dva ili četiri respektivno, pa je integralna nelinearnost daleko manja.

V. Zaključak

U radu je pokazano da koncept softverskog diterovanja izlaza iz anemometra sa nepokretnim ramovima nudi mogućnost znatno tačnijeg merenja energije ali i srednjeg smera vetra. U navedenoj literaturi je definisana i eksperimentalno proverena ideja merenja nelinearnim senzorom, kako sa jednim, tako i sa dva izlaza. U našem slučaju jedan izlaz predstavlja intenzitet (brzinu) vetra $|v|$, a drugi izlaz predstavlja smer (ugao) duvanja vetra φ.

Stoga velike nelinearnosti, vidljive sa kalibracionih dijagrama na Sl. 2. i Sl. 3. praktično nemaju uticaj na tačnost merenja energije odnosno srednjeg smera duvanja vetra - nih, pre svega, određuje tačnost primenjene kalibracione opreme i postupka. Kako je pokazano u radu [3] prilikom kalibrisanja anemometra sa nepokretnim ramovima korišćena je oprema najvišeg ranga – aerodinamički tunel.

Zahvalnica

Ovaj rad je podržan od strane Ministarstva prosvete, nauke i tehnološkog razvoja kroz institucionalno finansiranje naučno-istraživačkog rada na Fakultetu tehničkih nauka Univerziteta u Novom Sadu.

Literatura

[1] Vujicic, V., Ličina, B., Pejić, D., Sovilj, P. and Aleksandar, R., “Stochastic Measurement of Wind Power Using a Two-Bit A/D Converter”, Elsevier Measurement, Volume 152, February 2020, 107184, https://doi.org/10.1016/j.measurement.2019.107184

[2] Howard Austerlitz, Data Acquisition Techniques Using PCs, Book, Second Edition, 2003, DOI: https://doi.org/10.1016/B978-0-12-068377-2.X5000-1, ISBN: 978-0-12-068377-2

[3] L. Pugi et al., “Integrated Design and Testing of an Anemometer for Autonomous Sail Drones,”, J. Dyn. Sys. Meas. Control, vol. 140, no. 5, pp. 1–10, May.

[4] Platon Sovilj, Bojan Vujicic, Dragan Pejic, Aleksandar Radonjic, Vladimir Vujicic: “Stochastic Two-bit On-line Temperature Measurement with RTD Pt-100 Sensor Operating in a Nonlinear Mode”, 24th IMEKO TC4 International Symposium, 22nd International Workshop on ADC and DAC Modelling and Testing, IMEKO TC-4 2020, Palermo, Italy, September 14-16, 2020

[5] Vujicic V., Milovancev S., Pesaljevic M., Pejic D., Zupunski I.: “Low frequency stochastic true RMS instrument”, CPEM Digest (Conference on Precision Electromagnetic Measurements, Cat.No.98CH36254), 6-10 July 1998, Washington, DC, USA, DOI: 10.1109/CPEM.1998.699956

[6] Doktorska disertacija: Prilog optimizaciji performansi digitalnih merenja, Marjan Urekar – 2018, Fakultet tehničkih nauka, Univerzitet u Novom Sadu

[7] Attila Juhász, Vladimir Vujičić: “Contribution to Experimental Investigation of Fitting the Current Transformers’ Calibration Curves”, Journal on Processing and Energy in Agriculture, Biblid: 1821-4487 (2020) 24; 3-4, p 123-128, UDK: 621.314, DOI: 10.5937/jpea24-30719

[8] Doktorska disertacija: Metoda merenja snage i energije vetra zasnovana na merenju na intervalu, Boris Ličina – 2020, Fakultet tehničkih nauka, Univerzitet u Novom Sadu

Abstract

The paper analyzes the measurement of wind power and energy with an anemometer with fixed frames. Two scenarios are considered - measurement by standard sampling method (SSM) and measurement by two-bit stochastic digital measurement method (SDMM). A novelty is the software dithering of the anemometer output.

Measurement of wind power and energy with an anemometer without moving parts

Boris Ličina, Bojan Vujičić, Platon Sovilj, Vladimir Vujičić

Inženjerska indukcija – predlog definicije i jedna potvrda predloga

Bojan Vujičić, Boris Ličina, Platon Sovilj, Vladimir Vujičić

***Apstrakt* — U radu se predlaže definicija novog inženjerskog pojma nazvanog inženjerska indukcija. Predlog je primenjen u istraživanjima optimalne strukture uređaja za merenje snage i energije vetra i pokazao je svoju efikasnost i primenljivost.**

***Ključne reči* — matematička indukcija; deduktivni dokaz, induktivni dokaz; inženjerska indukcija.**

I. UVOD

PRE nego što izložimo osnovnu ideju ovog rada, važno je istaći da za matematičku idukciju nije "zaslužna" konkretna ličnost, niti je nastala konkretnog datuma. Ona je nastala kao posledica mnogobrojnih radova različitih matematičara, i ne samo matematičara, već i naučnika iz heterogenih naučnih disciplina. Najraniji tragovi matematičke indukcije sežu još u Euklidove *Elemente*, u treći vek p.n.e. Ovom prilikom se nećemo upuštati u dublje istorijske detalje njenog nastanka, ali ćemo spomenuti da se prva konkretna formulacija matematičke indukcije sreće kod Paskalovog rešenja aritmetičkog trougla 1654. godine.

Tekst koji sledi nema pretnziju da u bilo kom segmentu umanji značaj matematičke indukcije. Naprotiv, na ovom konkretnom primeru, princip "inženjerske" indukcije se pokazuje kao samo jedan od praktičnih načina za brže, a jednako korektno, rešenje problema koji se pred autore postavio. Tim pre, što se ponuđeno rešenje ima veoma jednostavnu potvrdu kroz hardversku realizaciju prikazanu u jednom od ranijih radova [1]. U ovom konkretnom slučaju prikazano je opšte rešenje "inženjerskom" indukcijom i opštim deduktivnim postupkom i na taj način izvršena potvrda tvrdnje na oba načina.

Ovaj rad nastao je iz inženjerske prakse, kao potreba da se na egzaktan, inženjerski i logički korektan način, ali lišen strogih pravila matematičke indukcije, na jednostavniji način reši jedan konkretan inženjerski problem. Opšte je poznata činjenica da su inženjeri u svojoj praksi često suočeni sa potrebom da nađu konkretno rešenje za svakodnevne izazove sa kojima se susreću u svom okruženju. Neretko, potreba za rešenjem je vremenski limitirana, pa se u takvim okolnostima pribegava različitim pristupima koji dovode do kvalitetnog i korektnog rešenja. Takav pristup ima svoje dobre i loše strane. Dobra stvar je što se brže dolazi do rešenja konkretnog problema, a loša stvar je što se uglavnom ne vodi računa o uopštavanju rešenja i njegovoj generalizaciji, o "široj slici".

Za razliku od ovog pristupa matematičari nastoje da objedine što veći broj problema u jedan opšti i da ponude njegovo rešenje, bez ulaženja u inženjersku suštinu da li je to nekome stvarno potrebno i hoće li ovakvo rešenje imati široku primenu u praksi. Imajući ovo u vidu, prilikom rešenja jednog konkretnog problema, nastao je ovaj rad kao težnja da "pomiri" ova dva, samo naizgled, oprečna pristupa. Da pokaže kako oba pristupa dovode do korektnog rešenja, da ne isključuju jedan drugog, već, naprotiv, da jedan drugog podupiru u zajedničkom cilju [2].

II. PROBLEM MERENJA SREDNJE VREDNOSTI PROIZVODA N NEPREKIDNIH SIGNALA

Pretpostavimo da se, tokom vremenskog intervala T, proizvod k signala meri korišćenjem dvobitne statističke digitalne merne metode (SDMM). U tom slučaju će izlazna vrednost množača koju u stvari generiše $(k-1)$ - binarni množač biti jednaka:

$$\overline{\Psi}(\mathrm{k}) = \frac{1}{\mathrm{N}} \cdot \sum_{\mathrm{i}=1}^{\mathrm{N}} \Psi_1(\mathrm{i}) \cdot \Psi_2(\mathrm{i}) \ldots \cdot \Psi_{\mathrm{k}}(\mathrm{i})$$
$$= \frac{1}{\mathrm{T}} \cdot \int_0^{\mathrm{T}} \mathrm{f}_1(\mathrm{t}) \cdot \mathrm{f}_2(\mathrm{t}) \ldots \cdot \mathrm{f}_{\mathrm{k}}(\mathrm{t}) \cdot \mathrm{dt} \qquad (1)$$

gde su $\Psi_1(\mathrm{i}), \Psi_2(\mathrm{i}), \ldots \Psi_{\mathrm{k}}(\mathrm{i})$ digitalizovane vrednosti k ulaznih signala i gde $\mathrm{N} \rightarrow \infty$ označava neograničen broj uzoraka (uzrokovanih neograničenom frekvencijom uzorkovanja) u vremenskom intrevalu T. Navedimo sada sledeću **teoremu:** Pretpostavimo da je $\sigma_{\bar{e}}^2(k) = \sigma_e^2(k)/N$ varijansa srednje greške merenja e proizvoda k signala merenih dvobitnom SDMM u vremenskom intervalu $[0, T]$. Ako sa $+g$ i $-g$ označimo naposke pragove odlučivanja dvobitnog AD konvertora tada, za konačnu frekvenciju uzorkovanja, pa time i ograničenu vrednost uzoraka N, možemo pisati:

Bojan Vujičić – Univerzitet u Novom Sadu, Fakultet tehničkih nauka Novi Sad, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: bojanvuj@uns.ac.rs).

Boris Ličina – Univerzitet Privredna Akademija Novi Sad, Fakultet za primenjeni menadžment, ekonomiju i finansije, Jevrejska 24/1, 11000 Beograd (e-mail: boris.licina@yahoo.com).

Platon Sovilj – Univerzitet u Novom Sadu, Fakultet tehničkih nauka Novi Sad, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: platon@uns.ac.rs).

Vladimir Vujičić – Univerzitet u Novom Sadu, Fakultet tehničkih nauka Novi Sad, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: vujicicv@uns.ac.rs).

$$\sigma_{\bar{e}}^2(k) = \frac{1}{N} \cdot \left[\frac{(2g)^k}{T} \cdot \int_0^T |f_1(t) \cdot f_2(t) \cdot \ldots \cdot f_k(t)| \cdot dt - \frac{1}{T} \cdot \int_0^T f_1^2(t) \cdot f_2^2(t) \cdot \ldots \cdot f_k^2(t) dt \right] \quad (2)$$

Dokaz: Umesto strogog i dugog deduktivnog dokaza, dajemo jednostavnu analizu (**u duhu matematičke indukcije**) koja potvrđuje opštu formulu (2).

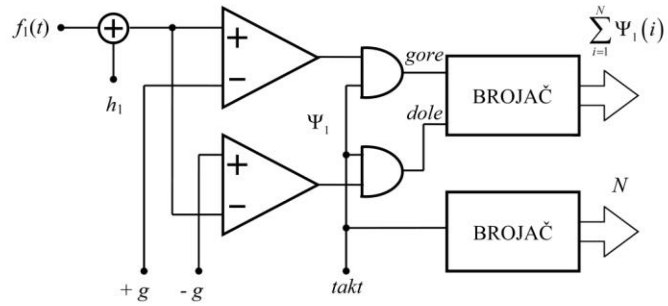
Korak 1: Za $k = 1$, izlazna vrednost je jednaka:

$$\Psi = f_1(t) + e \quad (3)$$

a varijansa srednje greške je izražena sa:

$$\sigma_{\bar{e}}^2(1) = \frac{1}{N} \cdot \left[\frac{2g}{T} \cdot \int_0^T |f_1(t)| \cdot dt - \frac{1}{T} \cdot \int_0^T f_1^2(t) \cdot dt \right] \quad (4)$$

Hardverska realizacija slučaja za $k = 1$, gde se ulaznom signalu $f_1(t)$ superponira diterski signal h_1, kako je to prikazano na Slici 1.



Sl. 1. Blok dijagram uređaja za merenje srednje vrednosti ulaznog signala.

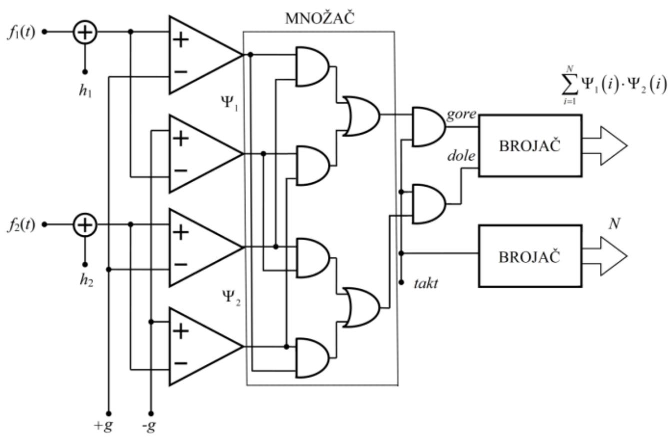
Korak 2: Za $k = 2$, izlazna vrednost je jednaka:

$$\Psi = f_1(t) \cdot f_2(t) + e \quad (5)$$

a varijansa srednje greške je izražena sa:

$$\sigma_{\bar{e}}^2(2) = \frac{1}{N} \cdot \left[\frac{(2g)^2}{T} \cdot \int_0^T |f_1(t) \cdot f_2(t)| \cdot dt - \frac{1}{T} \cdot \int_0^T f_1^2(t) \cdot f_2^2(t) \cdot dt \right] \quad (6)$$

Hardverska realizacija slučaja za $k = 2$, gde se ulaznim signalima $f_1(t)$ i $f_2(t)$ superponiraju međusobno nekorelisani diterski signali h_1 i h_2, respektivno, prikazana je na Slici 2.



Sl. 2. Blok dijagram uređaja za merenje proizvoda dva ulazna signala.

Korak 3: Primetimo da se izrazi (5) i (6) mogu zapisati kao:

$$\sigma_{\bar{e}}^2(1) = \frac{1}{N} \cdot \left[\frac{(2g)^0 \cdot 2g}{T} \cdot \int_0^T |1| \cdot |f_1(t)| \cdot dt - \frac{1}{T} \cdot \int_0^T [1^2 \cdot f_1^2(t)] \cdot dt \right] \quad (7)$$

i

$$\sigma_{\bar{e}}^2(2) = \frac{1}{N} \cdot \left[\frac{(2g)^1 \cdot 2g}{T} \cdot \int_0^T |f_1(t)| \cdot |f_2(t)| \cdot dt - \frac{1}{T} \cdot \int_0^T [f_1^2(t) \cdot f_2^2(t)] \cdot dt \right] \quad (8)$$

Pored toga, primetimo da, za bilo koji $k \geq 1$, prilikom dodavanja sledeće funkcije i sledećeg množača vredi:

a) prvi integral je pomnožen sa $2 \cdot g$,

b) funkcija unutar prvog integrala množi se sa apsolutnom vrednošću sledeće funkcije,

c) funkcija unutar drugog integrala množi se sa kvadratom sledeće funkcije.

Korak 4: Pretpostavimo da je teorema tačna za neko k, te da važe pretpostavke a), b) i c). Ostaje da pokažemo da važi i za $k + 1$.

$$\sigma_{\bar{e}}^2(k+1) = \frac{1}{N} \left[\frac{(2g)^k \cdot 2g}{T} \int_0^T |f_1(t) f_2(t) \ldots f_k(t)| \cdot |f_{k+1}(t)| \cdot dt - \frac{1}{T} \int_0^T [f_1^2(t) f_2^2(t) \ldots f_k^2(t)] \cdot f_{k+1}^2(t) dt \right] \quad (9)$$

Kad se navedene operacije izvrše, dobijamo:

$$\sigma_{\bar{e}}^2(k+1) = \frac{1}{N}\left[\frac{(2g)^{k+1}}{T}\int_0^T |f_1(t)\cdot f_2(t)\cdot \ldots \cdot f_{k+1}(t)|\cdot dt - \frac{1}{T}\int_0^T [f_1^2(t)\cdot f_2^2(t)\cdot \ldots \cdot f_{k+1}^2(t)]\cdot dt\right] \quad (10)$$

čime je teorema dokazana.

Primetimo da je dokaz “u duhu matematičke indukcije” ali da nije matematička indukcija. Autori su skloni da ovaj primenjeni postupak nazovu “inženjerska indukcija” i da ga u daljim istraživanjima detaljnije uopšte i formalizuju.

III. Deduktivna Potvrda

Dokažimo deduktivno formulu za varijansu srednje greške merenja proizvoda k signala dvobitnom SDMM. To ćemo uraditi u pet koraka, kao je to niže prikazano.

Korak 1: M_3 je konačan:

$$M_3 = \overline{(e-\bar{e})^3} \quad (11)$$

Može se strogo dokazati da važi $M_3 \le (2g)^{3k}$ odnosno da je treći centralni momenat slučajne greške e konačan (ograničen) i da to važi za svaki konačan prirodan broj k ($k \ge 1$).

Korak 2: Pošto je M_3 konačan [3] važi:

$$\sigma_{\bar{e}}{}^2 = \frac{\sigma_e^2}{N} \quad (12)$$

Korak 3: $\sigma_e^2 = ?$

Trenutna vrednost proizvoda k ulaznih signala je deterministička veličina a trenutna vrednost greške merenja e je slučajna veličina. Prema tome, one su međusobno statistički nezavisne pa je:

$$\begin{aligned} \Psi &= y_1\cdot y_2\cdot \ldots y_k + e \\ \sigma_\Psi^2 &= \sigma_{y_1y_2\cdots y_k}^2 + \sigma_e^2 \\ \sigma_e^2 &= \sigma_\Psi^2 - \sigma_{y_1y_2\cdots y_k}^2 \end{aligned} \quad (13)$$

Po definiciji je varijansa proizvoda k signala na vremenskom intervalu T data sa:

$$\sigma_{y_1y_2\cdots y_k}^2 = \frac{1}{T}\cdot\int_0^T f_1^2(t)\cdot f_2^2(t)\cdot \ldots \cdot f_k^2(t)\cdot dt - \left[\frac{1}{T}\cdot\int_0^T f_1(t)\cdot f_2(t)\cdot \ldots \cdot f_k(t)\cdot dt\right]^2 \quad (14)$$

Sa druge strane je:

$$\sigma_\Psi^2 = \overline{\Psi^2} - \overline{\Psi}^2 \quad (15)$$

Korak 4: Čemu je jednaka srednja vrednost izlaza fleš AD konvertora: $\overline{\Psi} = ?$ Zapišimo sledeći simbolički integral:

$$\overline{\Psi} = \int_{-(2g)^k}^{+(2g)^k} \Psi \cdot dP_\Psi \quad (16)$$

gde je:

$$\Psi = \Psi_1\cdot\Psi_2\cdot \ldots \Psi_k \quad (17)$$

a dP_Ψ elementarna verovatnoća da se Ψ desi:

$$dP_\Psi = dP_{y_1/t}\cdot dP_{y_2/t}\cdot \ldots \cdot dP_{y_k/t}\cdot dP_t\cdot dP_{h_1}\cdot dP_{h_2}\cdot \ldots \cdot dP_{h_k} \quad (18)$$

Kako je:

$$dP_{y_i/t} = \delta[y_i - f_i(t)]\cdot dy_i, \quad za \quad i = 1,2,\ldots k \quad (19)$$

$$dP_t = \frac{1}{t_2 - t_1}\cdot dt \quad (20)$$

$$dP_{h_i} = \frac{1}{2\cdot g}\cdot dh_i, \quad za \quad i = 1,2,\ldots k \quad (21)$$

konačno dobijamo:

$$\overline{\Psi} = \frac{1}{t_2-t_1}\cdot\int_{t_1}^{t_2} dt\cdot\int_{-2g}^{+2g}\delta[y_1 - f_1(t)]\cdot dy_1\cdot \int_{-2g}^{+2g}\delta[y_2 - f_2(t)]\cdot dy_2\cdot \ldots \cdot \int_{-2g}^{+2g}\delta[y_k - f_k(t)]\cdot dy_k\cdot \int_{-g}^{+g}\Psi_1\frac{dh_1}{2g}\cdot\int_{-g}^{+g}\Psi_2\frac{dh_2}{2g}\cdot \ldots \cdot \int_{-g}^{+g}\Psi_k\frac{dh_k}{2g} \quad (22)$$

Lako se pokazuje [4] da je za $y_i = \text{const.}$:

$$\int_{-g}^{+g}\Psi_i\frac{dh_i}{2g} = y_i, \quad za \quad i = 1,2,\ldots k \quad (23)$$

pa se gornji integral na kraju svodi na:

$$\overline{\Psi} = \frac{1}{t_2 - t_1} \cdot \int_{t_1}^{t_2} f_1(t) \cdot f_2(t) \cdot \ldots \cdot f_k(t) \cdot dt \qquad (24)$$

Korak 5: $\overline{\Psi^2} = ?$

Lako se pokazuje da u dvobitnoj SDMM generalno važi:

$$\Psi_i^l = (\text{sgn}\, \Psi_i)^l \cdot (2g)^{l-1} \cdot |\Psi_i| \qquad (25)$$

gde je *l* končan prirodan broj.

Za $l = 2$ što je naš slučaj $\Psi_i^2 = (2g) \cdot |\Psi_i|$, pa je:

$$\Psi^2 = \Psi_1^2 \cdot \Psi_2^2 \cdot \ldots \Psi_k^2 = \\ = (2g)^k \cdot |\Psi_1| \cdot |\Psi_2| \cdot \ldots |\Psi_k| = (2g)^k \cdot |\Psi| \qquad (26)$$

a srednja vrednost od Ψ^2 je:

$$\overline{\Psi^2} = (2g)^k \cdot \overline{|\Psi|} \qquad (27)$$

Ponavljajući gore navedenu proceduru, dobija se:

$$\overline{\Psi^2} = (2g)^k \cdot \frac{1}{t_2 - t_1} \cdot \int_{t_1}^{t_2} |f_1(t)| \cdot |f_2(t)| \cdot \ldots \cdot |f_k(t)| \cdot dt \qquad (28)$$

pa je najzad:

$$\sigma_e^2 = \frac{(2g)^k}{t_2 - t_1} \cdot \int_{t_1}^{t_2} |f_1(t) \cdot f_2(t) \cdot \ldots \cdot f_k(t)| \cdot dt \\ - \left[\frac{1}{t_2 - t_1} \int_{t_1}^{t_2} f_1(t) \cdot f_2(t) \cdot \ldots \cdot f_k(t) \cdot dt\right]^2 \\ - \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} [f_1^2(t) \cdot f_2^2(t) \cdot \ldots \cdot f_k^2(t)] \cdot dt \\ + \left[\frac{1}{t_2 - t_1} \int_{t_1}^{t_2} f_1(t) \cdot f_2(t) \cdot \ldots \cdot f_k(t) \cdot dt\right]^2 \qquad (29)$$

i konačno je:

$$\sigma_{\bar{e}}^2 = \frac{1}{N} \cdot \left\{\frac{(2g)^k}{t_2 - t_1} \cdot \int_{t_1}^{t_2} |f_1(t) \cdot f_2(t) \cdot \ldots \cdot f_k(t)| \cdot dt \right. \\ \left. - \frac{1}{t_2 - t_1} \cdot \int_{t_1}^{t_2} [f_1^2(t) \cdot f_2^2(t) \cdot \ldots \cdot f_k^2(t)] \cdot dt \right\} \qquad (30)$$

što je i trebalo dokazati.

IV. Zaključak

U radu su prikazana dva načina izvođenja opšte formule za grešku merenja srednje vrednosti proizvoda k neprekidnih analognih signala primenom SDMM. Jedan način je primenom novog induktivnog pristupa koji su autori nazvali inženjerska indukcija a drugi je generalni deduktivni način. Dobijena opšta formula u oba slučaja je identična i ima dalekosežne posledice i primenu. Jedna već ostvarena primena je u vrlo tačnom merenju energije vetra anemometrom sa šoljicama. Druga već ostvarena primena je u merenju temperature Pt100 senzorom u nelinearnom režimu.

Autori još traže opštu formalnu definiciju pojma inženjerske idukcije i otvoreni su za diskusiju i dopunu formalnog postupka primenjenog u ovom radu.

Zahvalnica

Ovaj rad je podržan od strane Ministarstva prosvete, nauke i tehnološkog razvoja kroz institucionalno finansiranje naučno-istraživačkog rada na Fakultetu tehničkih nauka Univerziteta u Novom Sadu.

Literatura

[1] Vujicic, V., Župunski, I., Milovancev, S. "Predetermination of the quantization error in digital measurement systems", IEEE Trans. on Instrumentation and Measurement, vol. 46, pp. 439-441, April 1997.

[2] Doktorska disertacija: Metoda merenja snage i energije vetra zasnovana na merenju na intervalu, Boris Ličina – 2020, Fakultet tehničkih nauka, Univerzitet u Novom Sadu

[3] V. Vranić, Probability and statistics. Zagreb: Tehnička knjiga. 1965. (in Croatian).

[4] M. Urekar et al., "Accuracy Improvement of the Stochastic Digital Electrical Energy Meter," Measurement, vol. 98, pp. 139-150, Feb. 2017.

Abstract

The paper proposes a definition of a new engineering concept called engineering induction. The proposal was applied in research of the optimal structure of devices for measuring wind power and energy and showed its efficiency and applicability.

Engineering induction - proposal of definition and one confirmation of proposal

Bojan Vujičić, Boris Ličina, Platon Sovilj, Vladimir Vujičić

Pregled doktorata u kojima je istraživana stohastička merna metoda

Dragan Pejić, Vladimir Vujičić

Apstrakt— **U radu je prikazan razvoj ideja vezanih za stohastičku adicionu analogno-digitalnu konverziju. Pregled je dat kroz prizmu doprinosa 18 doktorskih radova, počev od 1996. kada je odbranjen prvi doktorat na ovu temu. Neki radovi su u većoj meri bili teorijskog značaja, drugi su imali značajniju praktičnu težinu, ali su svi dali doprinos u razvoju i primeni stohastičke konverzije u poslednjih četvrt veka.**

Ključne reči—**Stohastički pristup; adaptivnost; merenja; diter.**

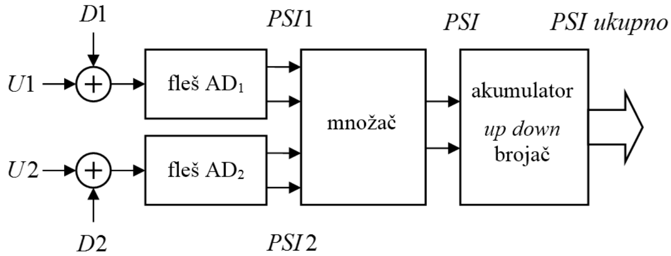
I. UVOD

Za početak primene stohastičkih principa, prvenstveno u takozvanom stohastičkom računanju, se navode [1] iz 1956. i [2] iz 1967. godine. U ova dva rada su postavljeni temelji konverzije analognih (napon) i digitalnih (broj u binarnom zapisu) vrednosti u povorku impulsa, tako da je informacija "utisnuta" u verovatnoću pojave jedinice. Pokazuje se da se, u slučaju operanada prikazanih na ovaj način, mnoge operacije obavljaju jednostavno: primenom jednostavnih kombinacionih ili sekvencijalnih mreža. Tako se množenje dva operanda u obliku povorke impulsa realizuje jednim AND logičkim kolom. Rezultat operacije je povorka impulsa kod koje je verovatnoća pojave jedinice približno jednaka proizvodu verovatnoća. Pored jednostavnosti izvođenja operacija, kao dobra osobina se navodi jednaka težina svih bita u povorci impulsa. Ovo za posledicu ima malu grešku u rezultatu, koja je posledica greške u bilo kom bitu operanda. Kod težinskih kodova je situacija druga: greška u bitu najveće vrednosti može da iznosi 50 %. Kao loše osobine ovog pristupa se navodi sporost, odnosno mala tačnost rezultata. Tačnost se može postići radom sa vrlo dugačkim operandima za šta je potrebna ili velika učestanost ili dugačko vreme računanja.

II. STOHASTIČKA ADICIONA ANALOGNO-DIGITALNA KONVERZIJA

Na Katedri za električna merenja Fakulteta tehničkih nauka u Novom Sadu, sredinom devedesetih godina XX veka, se počelo raditi na ideji koja je bazirana na primeni stohastike u merenjima. U osnovi ideje jeste primena analogno-digitalne (AD) konverzije fleš AD konvertorom male rezolucije, uz dodavanje šuma specijalnih osobina na ulazni naponski signal. Uloga aditivnog šuma jeste smanjivanje velike greške kvantizacije koja je sadržana u svakom rezultatu AD konverzije, što se postiže usrednjavanjem velikog broja susednih vrednosti.

Idejni tvorac ove ideje je profesor Vladimir Vujičić, koji je bio mentor prvog doktorskog rada [3] kolege Slobodana Milovančeva. U ovom doktorskom radu su postavljene osnove nove metode na bazi primene stohastike. Metoda je nazvana: stohastička adiciona AD konverzija. Reč je o analogno-digitalnoj konverziji jer se ulaznoj analognoj veličini na izlazu konvertora dodeljuje digitalna vrednost - broj. Termini "stohastička" i "adiciona" su upotrebljeni zbog primene specijalnog šuma (ditera) koji se dodaje na ulaznu analognu veličinu. U okviru doktorske teze je pokazana važna osobina ovakvog konverora - adaptivnost. Metrološke osobine rezultata konverzije zavise od broja osnovnih rezultata konverzije (N) nad kojima se vrši usrednjavanje. Pokazano je da preciznost rezultata konverzije raste sa kvadratnim korenom iz N. Na Sl. 1. je data blok šema stohastičkof adicionog AD konvertora sa dva generatora šuma.



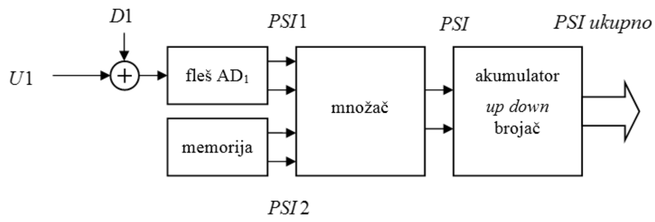
Sl. 1. Blok šema Stohastičkog adicionog analogno-digitalnog konvertora sa dva generatora ditera

U ovom radu je korišćen fleš AD konvertor sa svega tri moguća stanja na izlazu, kodovana vrednostima iz skupa {-1, 0, 1}. Operacija množenja se jednostavno izvodi upotrebom nekoliko logičkih kola i za rezultat se dobija vrednost iz polaznog skupa. Zbog ovoga se akumuliranje realizuje korišćenjem up-down brojača. Rezultat konverzije teži vrednosti određenog integrala proizvoda ulaznih napona. Ovakav blok je namenjen za određivanje aktivne energije i aktivne snage, u slučaju da su ulazni naponi srazmerni naponu i struji potrošača. Ako se na oba ulaza dovede napon, ovom metodom se može odrediti efektivna vrednost napona po definiciji.

Dragan Pejić, FTN, Univerzitet u Novom Sadu, Trg D. Obradovića 6, Novi Sad (email: pejicdra@uns.ac.rs)
Vladimir Vujičić, FTN, Univerzitet u Novom Sadu, Trg D. Obradovića 6 (email: vvujicic@uns.ac.rs)

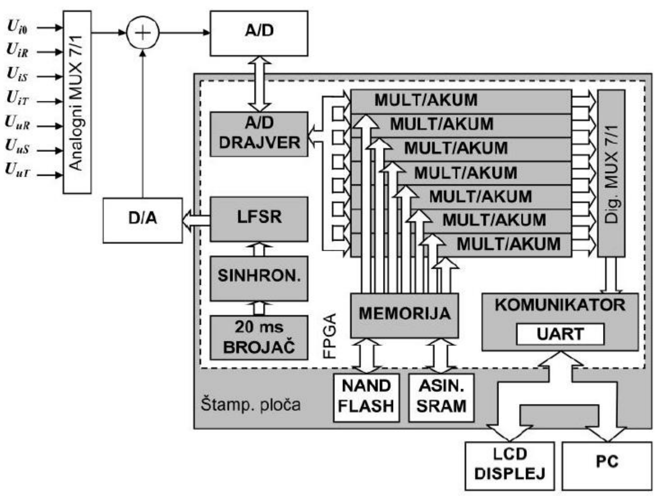
III. Frekventna analiza primenom stohastičke merne metode

U sledećem koraku se razmišljalo o primenjivosti stohastičke konverzije na određivanje koeficijenata razvoja u ortogonalni red. Ideja se nametnula sasvim logično budući da se koeficijenti razvoja u red određuju nalaženjem vrednosti određenog integrala proizvoda ulaznog napona i bazisne funkcije. Zaključeno je da nema potrebe realizovati bazisne funkcije u analognom obliku. Umesto toga, razmatrana je ideja prikazana na Sl. 2.



Sl. 2. Blok šema primene stohastičkog adicionog AD konvertora na određivanje jednog koeficijenta razvoja u red

Memorijski blok je iskorišćen za smeštanje dvobitnih odbiraka diterisane bazisne funkcije u trajanju jedne periode mrežnog napona i struje. Ovi odbirci se množe sa dvobitnim rezultatima fleš AD konverzije i rezultati se akumuliraju na potpuno isti način kao na Sl. 1. Ova tematika je teorijski i praktično analizirana u [4], gde je napravljen i korak dalje. Realizovan je sedmokanalni uređaj (tri napona i četiri struje), Sl. 3, kojim se vrši određivanje 50 harmonika po svakom od kanala.



Sl. 3. Blok šema sedmokanalnog uređaja za razvoj u red od 50 harmonika

Sve digitalne komponente (memorija, množači, akumulatori, itd...) su realizovane primenom FPGA (Field Programmable Gate Array) modula. Analogni deo sadrži sabirač i generisanje jednog diterskog signala. Na osnovu poznavanja harmonika napona i struja se određuju efektivne vrednosti, aktivne i reaktivne snage i energije. Uslov za funkcionisanje jeste bliskost učestanosti napona i struje nazivnoj mrežnoj učestanosti.

U [5] je izvršena generalizacija stohastičke metode i dat je detaljan matematički model. Analiziran je uticaj rezolucije primenjenog fleš AD konvertora za ulazni napon, kao i rezolucije diterisane bazisne funkcije. Pokazano je da je dovoljno da odbirci diterisanih bazisnih funkcija budu dva bita veće rezolucije nego odbirci analiziranog signala, kako ne bi dodatno povećavali grešku u određivanju razvoja u ortogonalni red. Takođe je pokazano da apsolutna greška određivanja koeficijenata razvoja u red ne zavisi od reda horamonika.

U [6] je prikazana primena stohastičke metode u određivanju harmonijskog sastava napona i struja kvara sa električnim lukom. Na osnovu dobijenih podataka je pokazano da se može efikasno utvrditi mesta nastanka kvara. Nad praktično dobijenim rezultatima je pokazano da se u vremenu od 50 ms do 60 ms može odrediti harmonijski sastav napona i struje sa greškom manjom od 2 %. Isti postupak sproveden u simulacionim uslovima daje mnogo manju grešku, reda 0.01 %.

Ofseta komparatora kojim se realizuje fleš AD konvertor je prepoznat kao jedan od najvećih uzročnika sistematske greške u okviru stohastičkog konvertora [7]. Realizovan je prototip merila, uz posebno obraćanje pažnje na smanjenje uticaja ofseta primenjenih komparatora. Pokazano je da se periodičnim obrtanjem ulaza komparatora, uticaj ofseta može smanjiti između 40 dB i 70 dB. Na ovaj način se ostvaruje mogućnost merenja aktivne energije u 15-minutnom intervalom sa greškom manjom od 100 ppm.

IV. Primena stohastičkog konvertora u biomedicinskim disciplinama

Nakon uspešne primene stohastičke metode na merenje parametara jednosmernog i prostoperiodičnog signala, složenoperiodičnog signala, u [8] je pokazana mogućnost primene u biomedicini: na merenje nestacionarnih EEG signala. Poznato je da je EEG signal male amplitude i zato lako prijemčiv za šum. Pokazano je da se primenom stohastičke digitalne metode može realizovati analiza EEG signala i što je najvažnije, postignuta je veća otpornost na prisustvo šuma nego kod klasičnih analognih i digitalnih metoda.

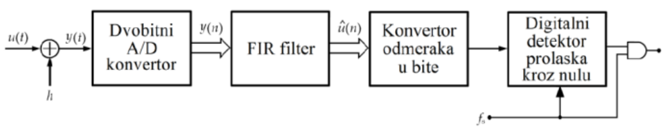
U [9] je pokazana mogućnost primene stohastičke digitalne metode u merenju elektrookulogrfskog signala. Primećeno je da dolazi do problema na krajevima mernog intervala usled takozvanog Gibsovog efekta. Problem je uspešno rešen preklapanjem vremenskih intervala. Preklapanjem se dobija rezerva, tako da se u kasnijoj analizi mogu odbaciti sami krajevi signala, upravo oni koji su problematični. Realizovan je sistem za testiranje, tako da se iz računara zadaje oblik signala na koji se dodaje diter i vrši fleš AD konverzija i konačno dobijeni odbirci vraćaju u računar. Na ovaj način je dobijen zatvoren sistem koji je korišćen za istraživanje primenjivosti stohastičke metode, uz ograničenje učestanosti generisanja ispitivanog signala, a time i učestanosti odabiranja fleš AD konvertorom, brzinom serijske komunikacije.

Primena stohastičke metode merenja na ERP (Event Related Potential) potencijalima je prikazana u [10]. Posmatrano je merenje latence ERP signala i vršne vrednosti P300 signala. Prilikom merenja latence ERP signala dobijena je greška do 1.5 %, dok je pri merenju vršne vrednosti P300 komponente dobijena greška u oko 13 %. I ovde je potvrđena ranije pomenuta otpornost na prisustvo šuma, što predstavlja prednost stohastičke metode nad drugim metodama.

V. GENERALIZACIJA STOHASTIČKE METODE

Doktorska teza [11] definiše pojam merenja na intervalu naspram standardnog pristupa koji je merenje u tački. Stohastička metoda je primer merenja na intervalu: rezultat stohastičke konverzije je vrednost određenog integrala proizvoda dve vremenski zavisne veličine. Kao što je već pokazano, mnoštvo veličina se određuje upravo na ovaj način: aktivna snaga i energija, efektivna vrednost, koeficijenti razvoja u ortogonalni red, itd... Samo usled razvoja tehnologije, očekuje se dalje poboljšanje metroloških performansi merenja na intervalu primenom stohastičke metode. Zahvaljujući adaptivnosti, sa porastom učestanosti odabiranja raste preciznost rezultata merenja dobijenih primenom stohastičkog pristupa.

Posebna pažnja tačnom određivanju mrežne učestanosti je posvećena u [12]. Tačno poznavanje vrednosti mrežne učestanosti je od presudnog značaja za kvalitet ostalih merenja u elektroenergetskom sistemu koja mogu biti obavljena primenom stohastičkog konvertora. Definisana je metoda za merenje mrežne učestanosti primenom stohastičke merne metode i FIR (Finite Impulse Response) filtera, Sl. 4.



Sl. 4. Blok šema za određivanje učestanosti napona primenom FIR filtera

Mala rezolucija primenjenog fleš AD konvertora, omogućava jednostavnu izvedbu FIR filtera bez korišćenja množača. Kako je moguća vrednost na izlazu fleš AD konvertora iz skupa -1, 0 i 1, umesto množenja vrednosti odbirka koeficijentom filtera, ovde se koeficijenti koriste sa promenjenim znakom, anuliraju se ili se koriste bez izmene, zavisno o vrednosti odbirka.

Ukoliko se poznaje tačna vrednost mrežne učestanosti, moguće je prevazići ograničenja na koja je ukazano u [4] i [5]. U [13] je obuhvaćen uticaj osobina diterskih signala, uticaj nepoklapanja trajanje merenja sa celim brojem perioda merenog signala, kao i uticaj odstupanja učestanosti osnovnog harmonika od nazivne vrednosti (50 Hz ili 60 Hz). Dat je niz preporuka za poboljšanje metroloških osobina merila koje se odnose na izvedbu pseudoslučajnih diterskih signala. Najvažniji doprinos disertacije je definisanje postupka kojim se u slučaju odstupanja učestanosti od nazivne vrednosti vrši određivanje vrednosti harmonika na osnovu skupa izmerenih vrednosti pseudoharmonika.

U radu [14] je prikazana je primena stohastičke merne metode u merenju napona i struje korišćenjem transformatora bez jezgra. Transformatori bez jezgra se odlikuju većom linearnošću, ali se retko koriste pošto na sekundaru daju napon srazmeran izvodu merene veličine. Primenom harmonijske analize bazirane na stohastičkoj mernoj metodi, ovaj problem se prevazilazi i što je još važnije, postaje prednost. Prirodno je da viši harmonici imaju manju vrednost, a zbog diferenciranja se dobija efekat pojačavanja n-tog harmonika faktorom n. U radu je dato rešenje za visokonaponski transformator bez jezgra, koji je opisan teorijski i simulaciono, a na prototipu sprovedena merenja su potvrdila prethodno dobijene rezultate.

Analiza uticaja variranja mrežne učestanosti na rezultate merenja primenom stohastičke metode u elektrodistributivnoj mreži je prikazana u [15]. Teorijskim, simulacionim i na posletku eksperimentalnim analizama potvrđena je mogućnost primene stohastičke metode sa 2-bitnim fleš A/D konvertorima u merenju, prevashodno osnovnog, a kasnije i viših harmonika u mreži, čak i u prisustvu značajne (7 puta veće od izmerene) gausovske varijacije mrežne frekvencije. Merne nesigurnosti rezultata simulacija merenja efektivne vrednosti osnovnog harmonika trougaonog (THD = 12 %) i testerastog signala (THD = 81 %) su ispod 0.006 %, čak i pri (za fleš A/D konvertore i SDMM metodu) frekvenciji odabiranja100 kHz.

Mogućnost primene stohastičke konverzije u nultim metodama je istraživana u [16]. Pokazano je teorijski, simulaciono i praktično da se zahvaljujući jednostavnoj strukturi stohastičkog konvertora ima mali broj izvora sistematske greške. Primenom metoda za smanjenje sistematskih grešaka na prihvatljivo male vrednosti, dobija se mogućnost vrlo tačnih i preciznih merenja malih napona. To kvalifikuje stohastičke metode za merenje u okolini nule.

U [17] je kombinovano korišćenje 4-bitnog fleš AD konvertora umesto dvobitnih. Dobijaju se 4-bitni odbirci ulaznih diterisanih veličina, što dalje zahteva komplikovaniji digitalni blok za množenje i akumuliranje. Sve digitalne funkcije su realizovane primenom FPGA modula. Dobitak je u višestruko boljim performansama koje se postižu u pogledu preciznosti rezultata merenja. Za postizanje dobre tačnosti merenja primenjene su poznate tehnike potiskivanja uticaja ofseta komparatora, kojih je sada 16 u fleš AD konvertoru, za razliku od svega dva komparatora u osnovnoj varijanti. Na vremenskom intervalu od 180 sekundi, na realizovanom prototipu je dobijena preciznost od 0.003 %, uz tačnost od 0.007 %.

U svim prethodnim praktičnim rešenjima diteri su generisani primenom LFSR (Linear Feedback shift Register) i digitalno-analognog (DA) konvertora. Pokazuje se da performanse DA konvertora ograničavaju kvalitet ditera u pogledu rezolucije i brzine generisanja, što se direktno odražava na finalne metrološke rezultate. U [18] je dat predlog za generisanje ditera bez upotrebe DA konverora. U radu je definisana metoda za generisanje šuma uniformne raspodele vrednosti odabiranjem trougaonog napona. Trenuci odabiranja trougaonog napona su određeni LFSR mehanizmom i nestabilnošću primenjene PLL petlje. Na ovaj način je izvršena randomizacija predvidivog (pseudoslučajnog) ponašanja LFSRa i nepredvidivog (u velikoj meri slučajnog) ponašanja PLL petlje i izbegnuta su ograničenja koja nameće primena DA konvertora.

Stohastička merna metoda je primenjena i optimizovana za merenje snage i energije vetra [19]. Ovo je sve važniji problem, obzirom na želju da se sve veći broj vetrogenratora instalira, ponekad na vrlo nedostupnim lokacijama. Jednostavnost stohastičke metode, pored malog broja izvora sistematskih grešaka čijim se otklanjanjem na dužem vremenskom intervalu postižu zapaženi metrološki rezultati, se karakteriše i malom potrošnjom električne energije. Ova činjenica dodatno kvalifikuje uređaj za merenje snage i energije vetra na bazi stohastičke metode za korišćenje u paralelnim merenjima na nedostupnim mestima koja se izvode radi ispitivanja potencijalnih lokacija na koje bi se postavljali vetrogeneratori.

Slično merenjima na intervalu, u radu [20] je predloženo upravljanje na intervalu: pristup koji se već uspešno pokazao u energetskim, biomedicinskim i drugim merenjima, ovde je proširen na fazi upravljanje. Rad prikazuje stohastički analogno-fazi konvertor (SAFC), uređaj koji radi sa podacima niske rezolucije u impulsnom režimu. Umesto rigorozne tačnosti, javljaju se slučajni događaji, zahvaljujući diterovanju nezanemarive greške kvantizacije po vrednosti pretvorene su u slučajnu grešku i kao takve proučavaju se statističkim metodama. Kombinovana je teorija verovatnoće sa fazi logikom, na način da se one međusobno dopunjuju.

VI. Zaključak

Stohastička adiciona analogno-digitalna konverzija je pojam po kojem je Katedra ze električna merenja Fakulteta tehničkih nauka u Novom Sadu prepoznatljiva. Metoda je primenjiva na određivanje aktivne energije, aktivne snage, efektivne vrednosti napona, za određivanje koeficijenata razvoja u red - svuda gde je potrebno odrediti vrednost određenog integrala proizvoda dva napona. Primenjivana je u elektrodistributivnim merenjima i biomedicini, kao i u fazi upravljanju.

Svaki doktorat na temu stohastičke merne metode je otkrio i ostavio neka otvorena pitanja i probleme. Na njih su odgovarali kasnije objavljivani članci ili sledeći doktorati. Neka od otvorenih pitanja koja su otvorena u poslednjih nekoliko doktorata su: a) merenje vrlo slabih, zašumljenih i izobličenih signala, b) istraživanje primene stohastičke digitalne merne metode optimalne rezolucije, c) realizacija stohastičkog digitalnog DFT procesora viokog reda, d) merenje nelinearnim senzorima, e) merenje posredno izmerenih nelinearnih veličina i f) merenje imitansnim senzorima.

Dalji razvoj stohastičke merne metode podrazumeva: a) prepoznavanje i otklanjanja izvora sistematske greške kako bi se tačnost metode i dalje povećavala i b) primenu sve bržih analognih i digitalnih komponenti, kako bi se u što kraćem vremenskom intervalu dobijali što precizniji rezultati merenja.

Literatura

[1] J. von Neumann, “Probabilistic logics and synthesis of reliable organisms from unreliable components,” in Automata Studies, Princeton University Press, 1956.

[2] B. R. Gaines, “Stochastic computing,” in American Federation of Information Processing Societies: Proceedings of the AFIPS ’67 Spring Joint Computer Conference, April 18-20, 1967, Atlantic City, New Jersey, USA, 1967.

[3] Slobodan Milovančev, "Adaptivni AD konvertor i njegova primena", FTN Novi Sad, 1996.

[4] Nebojša Pjevalica, "Merenja na elektrodistributivnoj mreži u frekvencijskom domenu", FTN Novi Sad, 2007.

[5] Velibor Pjevalica, "Stohastički digitalni procesor ortogonalnih funkcija", FTN Novi Sad, 2008.

[6] Branislav Santrač, "Precizno određivanje mesta kvara sa lukom u elektroenergetskom sistemu", Univerzitet u Kragujevcu, Tehnički fakultet Čačak, 2008.

[7] Dragan Pejić, "Stohastičko merenje električne snage i energije", FTN Novi Sad, 2010.

[8] Platon Sovilj, "Stohastičko digitalno merenje EEG signala", FTN Novi Sad, 2010.

[9] Jelena Đorđević Kozarov, "Metoda merenja elektrookulografskog signala ne intervalu sa preklapanjem vremenskih prozora", FTN Novi Sad, 2017.

[10] Milan Milovanović, "Metoda merenja moždanih ERP potencijala zasnovana na merenju harmonika epohe", FTN Novi Sad, 2105.

[11] Predrag Poljak, "Prilog razvoju i standardizaciji mernih metoda u merenjima u tački i merenjima na intervalu", FTN Novi Sad, 2012.

[12] Aleksandar Radonjić, "Merenja u frekvencijskom domenu u konceptu pametne mreže", FTN Novi Sad, 2013.

[13] Boris Antić, "Teorija realnog integrisanog merila harmonika", FTN Novi Sad, 2013.

[14] Dušan Čomić, "Stohastička metoda merenja napona i struje na visokom naponu", FTN Novi Sad, 2015.

[15] Željko Beljić, "Metoda merenja harmonika u realnoj distributivnoj mreži primenom AD konvertora dvobitne rezolucije", FTN Novi Sad, 2016.

[16] Bojan Vujičić, "Detekcija nule AD konvertorom niske rezolucije", FTN Novi Sad, 2017.

[17] Marjan Urekar, "Prilog optimizaciji digitalnih merenja", FTN Novi Sad, 2018.

[18] Nemanja Gazivoda, "Nova metoda za povećanje efektivne rezolucije stohastičkih mernih instrumenata visokih performansi", FTN Novi Sad, 2019.

[19] Boris Ličina, "Metoda merenja snage i energije vetra zasnovana na merenju na intervalu", FTN Novi Sad, 2020.

[20] Karolj Nađ, "Stohastički analogno fazi konvertor i njegova primena", FTN Novi Sad, 2013.

Abstract

The paper presents the development of ideas related to stochastic analog-to-digital conversion. The overview is given through the prism of the contribution of 18 PhD theses, starting from 1996, when the first doctorate on this topic was defended. Some works were more of theoretical significance, others had more significant practical weight, but all contributed to the development of stochastic conversions in the last 25 years.

Overview of PhD theses in which the stochastic measurement method was investigated

Dragan Pejić, Vladimir Vujičić

Primena mernog instrumenta VMP 20 za merenje snage u kolu naizmenične struje

Nemanja Vidović, Isidora Sabadoš, Atila Juhas i Saša Skoko

***Apstrakt*—U ovom radu predstavljen je jedan način primene mernog instrumenta VMP 20 u izvođenju laboratorijskih vežbi iz predmeta Električna merenja. Tema rada je merenje snage u kolu jednofazne struje za različite vrste opterećenja, predstavljanje rezultata merenja kao i mogućnost praktičnog izvođenja laboratorijske vežbe u realnom vremenu. Pažljivom organizacijom laboratorijskih vežbi, kao što je prikazano u ovom radu, moguće je uspostaviti maksimalnu korelaciju između osnovne teorije električnih merenja, mernih metoda kao i drugih stručnih predmeta kao što su Osnove elektrotehnike i teorija naizmeničnih struja.**

***Ključne reči*—merenje snage, merne metode, laboratorijske vežbe.**

I. Uvod

Savremeni pravci u obrazovnom sistemu naglašavaju razvijanje što ve'eg broja praktičnih veština budućih tehničara i inženjera. U tu svrhu potrebno je obezbediti odgovarajući broj stručnih predmeta u čijem sastavu će biti predviđen odgovarajući laboratorijski rad. Laboratorijski rad je veoma važan, jer se u okviru njega neprestano prožimaju teorijski principi i praktične realizacije u okviru neke stručne discipline.

Veoma je važno naglasiti potrebu primene savremenih mernih uređaja i mernih sistema u izvođenju laboratorijskih vežbi. Na taj način se stvaraju dobri preduslovi za maksimalno iskorišćenje potencijala izvođenja odgovarajuće laboratorijske vežbe.

U ovom radu prikazana je primena mernog instrumenta VMP 20 u izvođenju laboratorijske vežbe iz oblasti merenja snage u kolu jednofazne struje. Instrument pruža izuzetne pogodnosti u radu, jednostavno priključenje u električno kolo, jednostavnu upotrebu. Pored toga na prednjem displeju instrumenta prikazuju se osnovne merene veličine, ali se najveća vrednost ogleda u tome što za instrument postoji odgovarajući softver VMPCalc. Na taj način može se meriti više veličina nego što se prikazuje na displeju instrumenta. Uz primenu odgovarajućeg softvera instrument VMP 20 pruža mogućnost merenja na daljinu i razmene osnovnog prozora softvera za merenje preko video linka. Ovakvom organizacijom laboratorijskog rada, pruža se mogućnost primene mernog instrumenta VMP 20 u organizaciji onlajn laboratorijskih vežbi u okviru udaljenih laboratorija.

Nemanja Vidović – Elektrotehnička škola "Mihajlo Pupin", Futoška 17, 21000 Novi Sad, Srbija (e-mail: vidovicnemanja123@gmail.com).

Isidora Sabadoš – Elektrotehnička škola "Mihajlo Pupin", Futoška 17, 21000 Novi Sad, Srbija (e-mail: lapislazuli.ns@gmail.com).

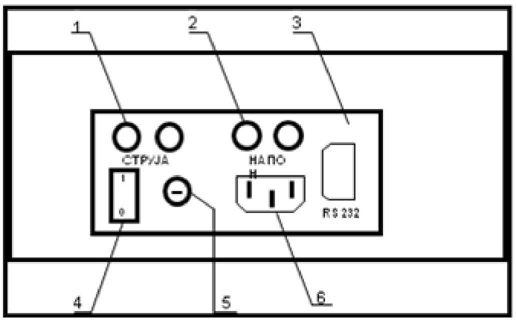
Atila Juhas, Fakultet tehničkih nauka Novi Sad, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija, (e-mail: juhiatti@gmail.com)

Saša Skoko – Elektrotehnička škola "Mihajlo Pupin", Futoška 17, 21000 Novi Sad, Srbija, (e-mail: sasaskoko78@gmail.com).

II. Merni instrument VMP 20

Meni instrument VMP 20 razvijen je na Fakultetu tehničkih nauka u Novom Sadu. Na Sl. 1. prikazan je izgled mernog instrumenta, a na Sl. 2. prikazana je priključna ploča.

Sl. 1. Izgled mernog instrumenta VMP 20.

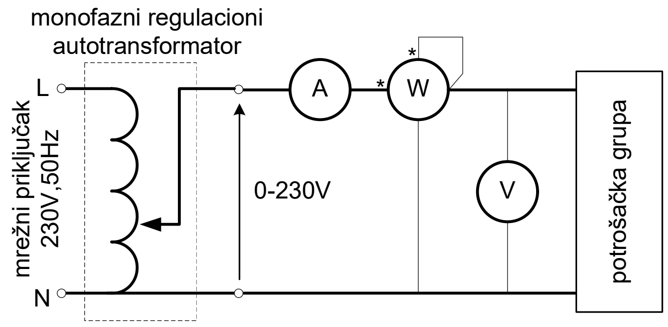


Sl. 2. Izgled priključne ploče instrumenta VMP 20.

Na Sl. 2. brojevima su označeni sledeći priključni kontakti: 1- strujni priključni kontakti, 2 - naponski priključni kontakti, 3- konektor za povezivanje sa računarom sa RS232 komunikacionim protokolom, 4 - prekidač za uključenje i isključenje mernog instrumenta, 5- osigurač, 6 - konektor za priključenje kabla za napajanje mernog sistema. Rad mernog sistema instrumenta zasniva se na stohastičkoj adicionoj A/D konverziji (SAADK). Metoda omogućava nov pristup merenjima na električnoj mreži, nezavisna je od tehnološkog razvoja i elektronskih komponenata, zahteva jednostavan hardver i adaptivnu preciznost merenja [1-4].

III. Merenje snage u kolu naizmenične struje

Standardni pristup izvođenju laboratorijskih vežbi podrazumeva primenu standardnih analognih mernih instrumenata. Na Sl. 3. prikazana je principska električna šema za izvođenje laboratorijske vežbe za merenje snage u kolu naizmenične struje primenom ampermetra, voltmetra i vatmetra [5].



Sl. 3. Električna šema za merenje snage u kolu jednofazne struje primenom standardnih analognih mernig instrumenata.

Na osnovu prikazane merne metode sa Sl. 3. napon, struja i aktivna snaga određuju se direktno očitavanjem mernih instrumenata, dok se prividna snaga, reaktivna snaga i faktor snage određuju indirektno primenom proračuna.

$$S = U \cdot I \tag{1}$$

$$cos\,\varphi = \frac{P}{S} \tag{2}$$

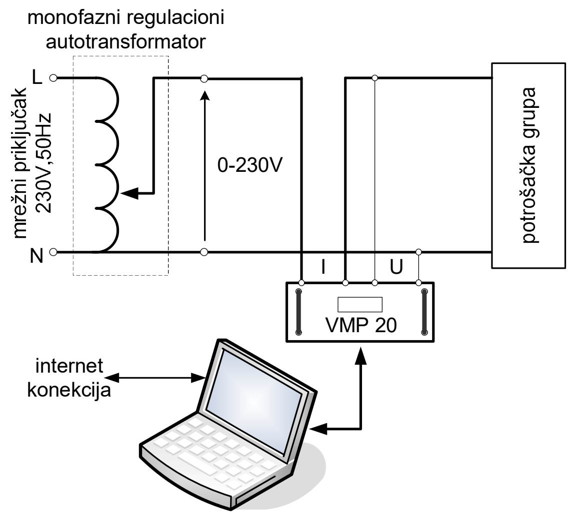
$$Q = \sqrt{S^2 - P^2} \tag{3}$$

Pri čemu u (1), (2), (3), veličine imaju sledeće značenje: I - izmerena struja, U - izmereni napon, P - izmerena aktivna snaga, S - prividna snaga, $cos\varphi$ - faktor snage, Q - reaktivna snaga. Na sl. 4. prikazan je izgled postavljene laboratorijske vežbe za merenje snage u kolu jednofazne struje.

Sl. 4. Izgled merne opreme za izvođenje laboratorijske vežbe primenom standardnih analognih mernih instrumenata.

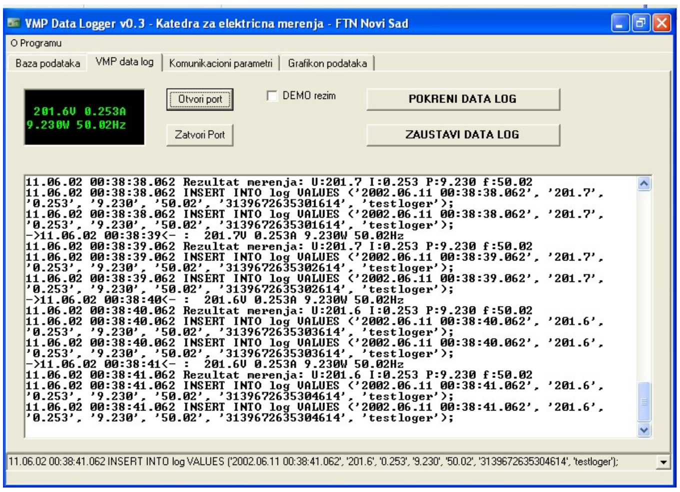
Osnovni nedostaci ovako postavljene laboratorijske vežbe ogledaju se u sledećem: mere se samo tri veličine (napon, struja i aktivna snaga) dok se ostale veličine izračunavaju, merni postupak relativno dugo traje zbog očitavanja analognih mernih instrumenata, ne postoji mogućnost izvođenja laboratorijske vežbe na daljinu.

Za razliku od standardnih mernih metoda koje se zasnivaju na primeni analognih mernih instrumenata, primena instrumenta VMP 20 pruža mnogo više mogućnosti. Na Sl. 5. prikazana je principska šema primene mernog instrumenta VMP 20 u merenju snage u kolu naizmenične struje.



Sl. 5. Električna šema za merenje snage u kolu jednofazne struje primenom mernog instrumenta VMP 20.

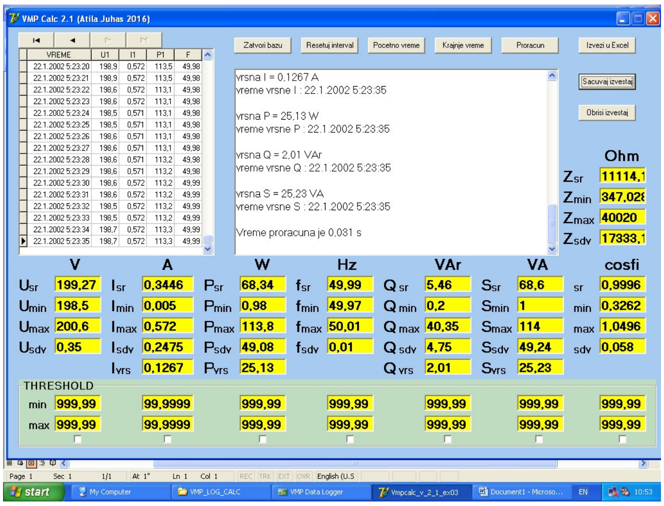
Primenom odgovarajućeg softvera VMPCalc moguće je u realnom vremenu meriti sledeće veličine: napon, struju, aktivnu snagu i frekvenciju. Na Sl. 6. prikazan je izgled osnovnog prozora softvera VMPCalc.



Sl. 6. Izgled osnovnog prozora softvera VMPCalc.

Na Sl. 6. možemo uočiti displej u gornjem desnom uglu prozora na kome se u realnom vremenu prikazuju vrednosti napona, struje aktivne snage i frekvencije potrošačke grupe. Ovaj način prikazaivanja izuzetno je pogodan za praćenje vrednosti merenih veličina u realnom vremenu ako je računar povezan na internet i postoji mogućnost deljenja ekrana računara putem odgovarajućeg video linka.

Pored prikazanog prozora na Sl. 6 softver za obradu podataka VMPCalc pruža mogućnost merenja više veličina na definisanom intervalu. Na Sl. 7. prikazan je izgled prozora softvera VMPCalc u kome se prikazuju sve veličine koje može da meri instrument VMP 20.



Sl. 7. Izgled prozora softvera VMPCalc prikaz više veličina merenih na intervalu.

U Tabeli I pregledno su prikazane električne veličine koje se mogu meriti instrumentom VMP 20, na definisanom intervalu.

TABELA I
PRIKAZ VELIČINA KOJE INSTRUMENT MERI NA INTERVALU

Naziv veličine	Oznaka i jedinica
Napon	*U* [V]
Struja	*I* [A]
Aktivna snaga	*P* [W]
Frekvencija	*f* [Hz]
Reaktivna snaga	*Q* [VAr]
Prividna snaga	*S* [VA]
Faktor snage	*cosφ* [-]
Impedansa potrošača	Z [Ω]

IV. PRIMENA INSTRUMENTA VMP 20 I REZULTATI MERENJA

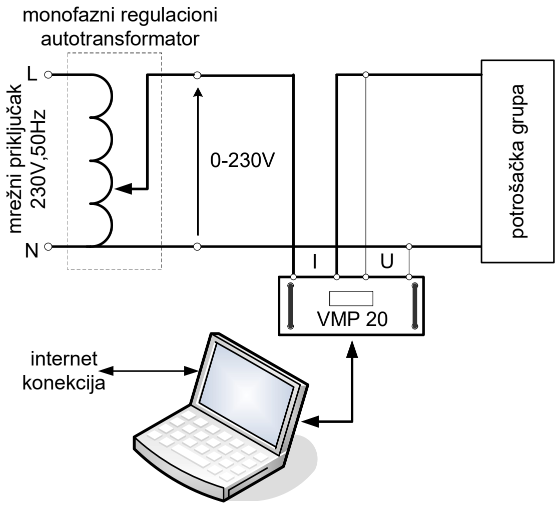
Laboratorijka vežba je tako koncipirana da je jasno definisan zadatak, data je električna šema i spisak ogledne opreme i kratko uputstvo za realizaciju laboratorijske vežbe. Učenik ili student, pre dolaska u laboratoriju, potrebno je da prouči tekst postavke laboratorijske vežbe i pripremiti se za nju.

U slučaju da se vežba radi u realnom vremenu, odnosno, onlajn, učenik ili student treba da bude pored svog računara u zakazanoj satnici. Neophodan preduslov je da postoji pristup internetu u laboratoriji za električna merenja.

Za rad u realnom vremenu preko interneta nastavnik ili saradnik su moderatori i definišu dinamiku izvođenja laboratorijske vežbe, dok učenik ili student preko video linka prate tok izvođenja laboratorijske vežbe.

U toku neposrednog merenja prozor softvera VMPCalc dostupan je putem video linka studentima ili učenicima i samostalno mogu izvršiti očitavanje merenih veličina.

Principski izgled električne šeme ispitne stanice za merenje snage u kolu naizmenične struje primenom instrumenta VMP20 u realnom vremenu prikazan je na Sl. 8.



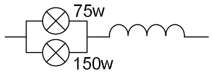
Sl. 8. Principski izgled električne šeme za merenje snage u kolu jednofazne struje primenom instrumenta VMP20.

Na Sl. 9. prikazan je izgled ispitne stanice u laboratoriji sa povezanom mernom opremom i odgovarajućim računarom.

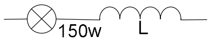
Sl. 9. Izgled ispitne stanice u laboratoriji sa povezanom mernom opremom i odgovarajućim računarom.

Tok izvođenja laboratorijske vežbe ogleda se u tome da se instrumentom VMP 20 mere električne veličine potrošača za različite kombinacije u potrošačkoj grupi. U Tabeli II prikazani su rezultati merenja za različite kombinacije elemenata potrošačke grupe.

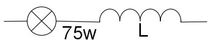
TABELA II
TABELARNI PRIKAZ REZULTATA MERENJA



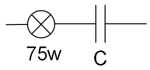
1. kombinacija		
U	200.8	[V]
I	0.828	[A]
P	141.5	[W]
S	166.26	[VA]
*cos*φ	0.853	
Q	86.45	[VAr]



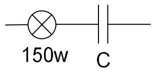
2. kombinacija		
U	201.4	[V]
I	0.578	[A]
P	108.7	[W]
S	116.41	[VA]
*cos*φ	0.932	
Q	41.91	[VAr]



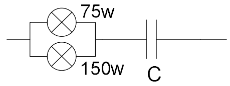
3. kombinacija		
U	202.9	[V]
I	0.311	[A]
P	61.85	[W]
S	63.101	[VA]
*cos*φ	0.982	
Q	11.36	[VAr]



4. kombinacija		
U	203.4	[V]
I	0.221	[A]
P	22.72	[W]
S	44.95	[VA]
*cos*φ	0.503	
Q	38.66	[VAr]



5. kombinacija		
U	203.6	[V]
I	0.256	[A]
P	9.452	[W]
S	52.12	[VA]
*cos*φ	0.181	
Q	51.07	[VAr]



6. kombinacija		
U	203.8	[V]
I	0.262	[A]
P	3.029	[W]
S	53.39	[VA]
*cos*φ	0.057	
Q	52.86	[VAr]

Završni deo vežbe učenik ili student radi samostalno. On se ogleda u korelaciji između teorije naučene iz Osnova elektrotehnike koje je neophodno povezati sa rezultatima dobijenim u laboratoriji za električna merenja.

Naime, neophodno je za odgovarajuće kombinacije potrošačke grupe konstruisati u odgovarajućoj razmeri trouglove snaga (aktivna, reaktivna, prividna) i dijagrame dobijene konstrukcionom metodom uporediti sa znanjima usvojenim na predmetu Osnove elektrotehnike. Proveriti da li komponente aktivne reaktivne i prividne snage čine pravougli trougao i eventualna odstupanja komentarisati stručno. Na taj način uspostavlja se snažna korelacija između osnovne teorije električnih merenja i teorije osnova elektrotehnike odnosno praktičnim merenjima proverava se da li važe osnovni teorijski principi u elektrotehnici.

V. ZAKLJUČAK

Cilj ovog rada je da približi jedan način izvođenja laboratorijskih vežbi iz električnih merenja. Obuhvaćena je jedna veoma važna tema a to je merenje snage u kolu naizmenične struje. Posebnu važnost u radu predstavlja primena savremenog mernog instrumenta VMP 20 i odgovarajućeg softvera VMPCalc za obradu merenih podataka. Zahvaljujući primeni softvera u radu je posebno akcentovana mogućnost primene instrumenta za izvođenje laboratorijskih vežbi u realnom vremenu i putem onlajn vežbi, odnosno video-linka. Na taj način podiže se kvalitet realizacije laboratorijskih vežbi, ali i ono što je još važnije, pruža se mogućnost pristupanja, praćenja i neposrednog rada učenika ili studenata koji nisu u mogućnosti iz objektivnih razloga da prisustvuju neposrednom izvođenju laboratorijskih vežbi. Iskustvo stečeno u radu sa instrumentom VMP20 je izuzetno pozitivno pa na ovom mestu želimo da istaknemo podršku autorima instrumenta na daljem hardverskom i softverskom usavršavanju mernog instrumenta.

ZAHVALNICA

Na ovom mestu želimo da se zahvalimo Vladimiru Vujičiću profesoru Fakulteta tehničkih nauka u penziji na velikoj podršci i savetima o mogućnostima primene instrumenta u neposrednoj laboratorijskoj praksi kao i pravcima primene u ravoju laboratorijskih merenja na daljinu i onlajn organizaciji laboratorijskih vežbi.

LITERATURA

[1] A. Juhas, "Hardverska i softverska podrška trofaznim merenjima snage i energije u ED mreži", master rad, Katedra za električna merenja, Fakultet tehničkih nauka u Novom Sadu, Novi Sad, Srbija, 2017.

[2] V. Vujičić, I. Župunski, Z. Mitrović, M. Sokola, "Merenje u tački i intervalu", Proc. of IMECO XIX, World Congress, Lisabon, Portugal, , pp. 1128-1132, No. 480, septembar, 2009.

[3] V. Vujičić, S. Milovančev, "Digitalni instrument za merenje srednje vrednosti proizvoda dava analogna periodična signala", YU patent, broj 48195, 1995.

[4] V. Pjevalica, V. Vujičić, "Further Generation of Low-Frequency True-RMS Instrument", IMTC, pp. 1008-1011, May, 2005.

[5] P. Dobanovački, S. Skoko, "Električna merenja-praktikum laboratorijskih vežbi", Elektrotehnička škola "Mihajlo Pupin", Novi Sad, Srbija, 2016.

ABSTRACT

This paper presents a way of applying the measuring instrument VMP 20 in performing laboratory exercises in the subject of Electrical Measurements. The topic of the paper is the measurement of power in a single-phase current circuit for different types of loads, the presentation of measurement results as well as the possibility of practical laboratory exercises in real time. By careful organization of laboratory exercises, as shown in this paper, it is possible to establish the maximum correlation between the basic theory of electrical measurements, measurement methods and other professional subjects such as Fundamentals of Electrical Engineering and Theory of Alternating Currents.

Practical implementation of the VMP 20 for measuring power in an alternating current circuit

Nemanja Vidović, Isidora Sabadoš, Atila Juhas i Saša Skoko

Primena mernog instrumenta VMP 20 za izvođenje laboratorijske vežbe - popravka faktora snage

Isidora Sabadoš, Nemanja Vidović, Atila Juhas i Saša Skoko

***Apstrakt*—U ovom radu predstavljen je jedan način primene mernog instrumenta VMP 20 u izvođenju laboratorijskih vežbi iz predmeta Električna merenja. Tema rada je popravka faktora snage u jednofaznom sistemu napajanja. Radom je obuhvaćen osnovni teorijski princip kompenzacije reaktivne snage i dat je prikaz izvođenja laboratorijske vežbe. Primena instrumenta VMP 20 i odgovarajućeg softvera VMPCalc ima posebnu pogodnost u organizaciji laboratorijskih vežbi na daljinu- online, što je u radu posebno naznačeno.**

***Ključne reči*—popravka faktora snage, VMP20, laboratorijske vežbe, online-laboratorija.**

I. UVOD

Savremeni pristup stručnom obrazovanju tehničara i inženjera podrazumeva povećani pristup razvoju praktičnih veština. Zato je neophodno obezbediti u obrazovnom procesu dovoljno praktičnog rada u okviru odgovarajućih laboratorijskih vežbi kako bi se praktično demonstrirao i proverio što veći broj teorijskih principa.

Zato organizacija laboratorijskih vežbi iz Električnih merenja ima poseban značaj. U okviru planiranja i realizacije vežbi potrebno je postaviti (osmisliti) odgovarajući broj zadataka koji bi uspešno i smisleno povezali osnovne teorijske principe i njihove praktične aspekte.

U ovom radu prezentovan je jedan takav pristup. Naime, nastojali smo da kroz organizaciju jedne laboratorijske vežbe izvršimo sintezu osnovnih teorijskih principa iz teorije Električnih merenja i da uspostavimo korelaciju sa osnovnim teorijskim sadržajima iz predmeta Osnove elektrotehnike 2 - teorija naizmeničnih struja, odnosno, kompenzacija reaktivne snage (popravka faktora snage). Prilikom postavke i organizacije laboratorijske vežbe korišćen je merni instrument VMP20 i njegov aplikativni softver VMPCalc za prikazivanje rezultata merenja. S obzirom da primena softvera VMPCalc omogućava rad u realnom vremenu, koncept realizacije laboratorijske vežbe je da se ukaže na mogućnosti primene instrumenta VMP20 u organizaciji i izvođenju laboratorijskih vežbi na daljinu (online laboratorijske vežbe, udaljena laboratorija). Na taj način se daje poseban doprinos podizanju kvaliteta organizacije i realizacije laboratorijskih vežbi.

Isidora Sabadoš – Elektrotehnička škola "Mihajlo Pupin", Futoška 17, 21000 Novi Sad, Srbija (e-mail: lapislazuli.ns@gmail.com).

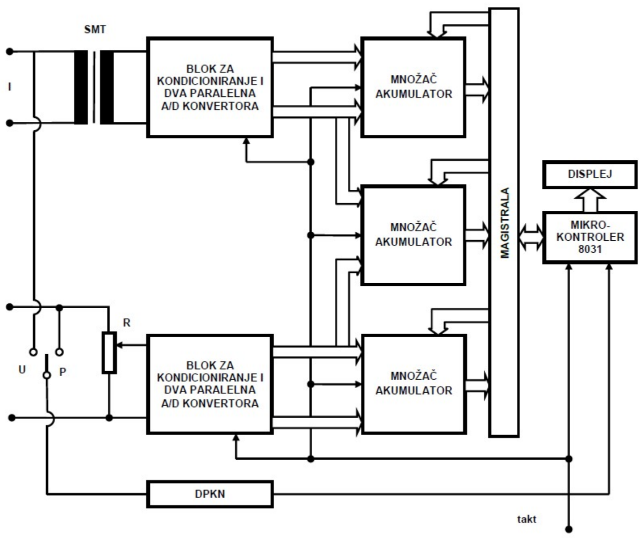
Nemanja Vidović – Elektrotehnička škola "Mihajlo Pupin", Futoška 17, 21000 Novi Sad, Srbija (e-mail: vidovicnemanja123@gmail.com).

Atila Juhas, Fakultet tehničkih nauka Novi Sad, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija, (e-mail: juhiatti@gmail.com)

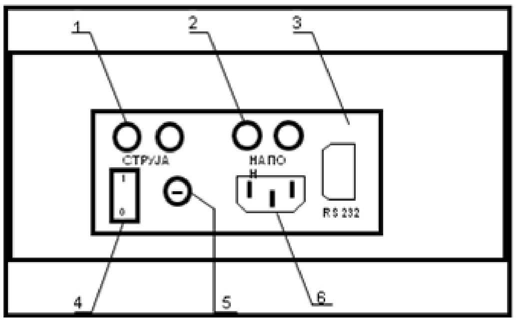
Saša Skoko – Elektrotehnička škola "Mihajlo Pupin", Futoška 17, 21000 Novi Sad, Srbija, (e-mail: sasaskoko78@gmail.com).

II. MERNI INSTRUMENT VMP 20

Meni instrument VMP 20 razvijen je na Fakultetu tehničkih nauka u Novom Sadu [1]. Na Sl. 1. prikazan je izgled osnovne blok šeme mernog instrumenta, a na Sl. 2. prikazana je priključna ploča.



Sl. 1. Izgled osnovne blok šeme mernog instrumenta VMP 20.



Sl. 2. Izgled priključne ploče instrumenta VMP 20.

Na Sl. 2. brojevima su označeni sledeći priključni kontakti: 1- strujni priključni kontakti, 2 - naponski priključni kontakti, 3 - konektor za povezivanje sa računarom sa RS232 komunikacionim protokolom, 4 - prekidač za uključenje i isključenje mernog instrumenta, 5- osigurač, 6 - konektor za priključenje kabla za napajanje mernog sistema. Rad mernog sistema instrumenta zasniva se na stohastičkoj adicionoj A/D konverziji (SAADK) [2],[3],[4]. Metoda omogućava nov pristup merenjima na električnoj mreži, nezavisna je od tehnološkog razvoja i elektronskih komponenata, zahteva jednostavan hardver i adaptivnu preciznost merenja.

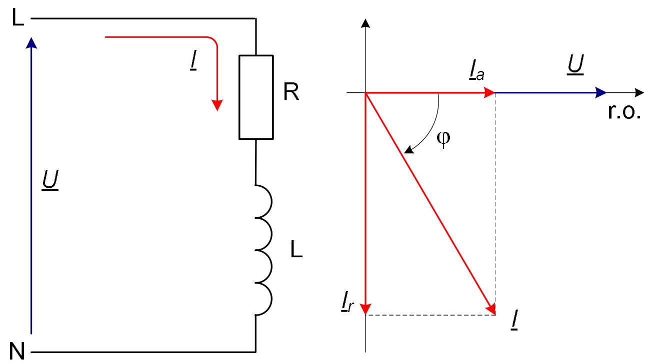
III. Osnovni principi kompenzacije reaktivne snage

Faktor snage u električnoj mreži zavisi od karaktera priključenih prijemnika. Ukoliko se karakter prijemnika tokom vremena menja, to znači da će se menjati i faktor snage. Faktor snage definiše se kao odnos aktivne i prividne snage i dat je izrazom (1) [5].

$$cos\,\varphi = \frac{P}{S} \qquad (1)$$

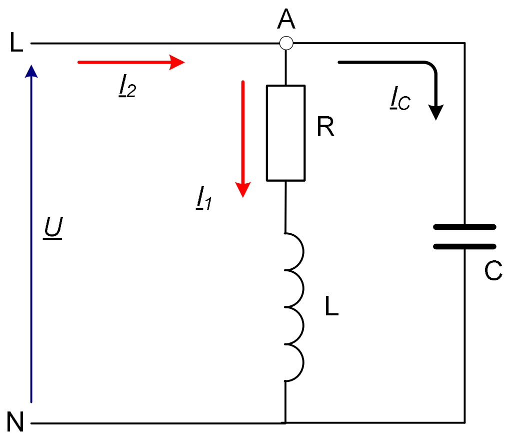
Ako je opterećenje električne mreže čisto omsko tada je faktor snage jednak 1. Ako je opterećenje u električnoj mreži termogeno-induktivnog karaktera tada se faktor snage kreće u intervalu od 0 - 1 u zavisnosti od toga koliki je udeo termogenog i induktivnog opterećenja u potrošačkoj grupi. Propisima je predviđeno da faktor snage u električnoj mreži bude što bliži 1, odnosno da bude iznad 0.95. S obzirom da su u proizvodnim pogonima prijemnici najčešće elektromotori, transformatori i slični omsko-induktivni prijemnici, oni predstavljaju značajne potrošače reaktivne snage pa je najčešće kod ovakvih potrošačkih grupa faktor snage ispod 0.95. Zato je u takvim slučajevima neophodno izvršiti kompenzaciju reaktivne snage i popravku faktora snage na zadovoljavajuću vrednost.

Na Sl. 3. prikazan je jedan RL potrošač povezan u kolo jednofazne struje sa odgovarajućim fazorskim dijagramima napona i struja.



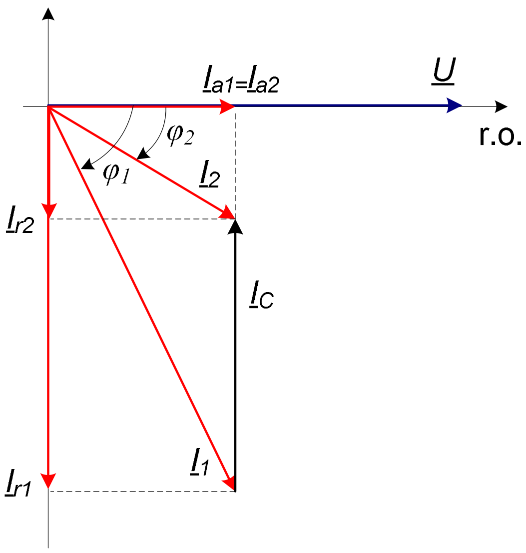
Sl. 3. RL potrošač povezan u kolo jednofazne struje sa odgovarajućim fazorskim dijagramom napona i struja.

Na Sl. 3. oznake imaju sledeće značenje: R - omski potrošač, L - induktivni potrošač, $\underline{U}$ - fazor napona na potrošaču, $\underline{I}$- fazor struje potrošača, $\underline{I}_a$ -fazor aktivne komponente struje potrošača, $\underline{I}_r$ -fazor reaktivne komponente struje potrošača, φ- fazni pomeraj između napona i struje potrošača. Pod uslovom da je vrednost napona u električnom kolu konstanta i da se tokom vremena ne menja, vrednost reaktivne snage zavisi od reaktivne komponente struje u električnom kolu. Što je potrošač dominantnijeg induktivnog karaktera to je veća reaktivna komponenta struje. Prema tome, ako želimo da kompenzujemo reaktivnu snagu potrošača, potrebno je na određeni način kompenzovati reaktivnu komponentu struje potrošača. Kompenzacija reaktivne komponente struje vrši se paralelnim vezivanjem kondenzatora sa RL potrošačem kao što je prikazano na Sl. 4.



Sl. 4. Kondenzator C paraleno povezan RL potrošaču u cilju kompenzacije reaktivne snage (struje).

Na Sl. 5. prikazan je fazorski dijagram napona i struja u električnom kolu pre i posle povezivanja kondenzatora.

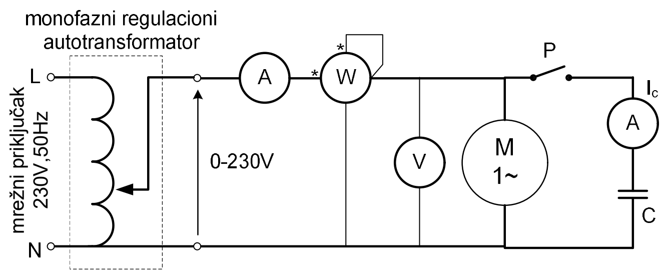


Sl. 5. Fazorski dijagram napona i struja električnog kola pre i posle povezivanja kondenzatora.

Pri čemu je značenje veličina na Sl. 5. sledeće: $\underline{I}_C$ - fazor struje kondenzatora, sa indeksom 1 označene su struje i njihove komponente u električnom kolu pre povezivanja kondenzatora, a sa indeksom 2 označene su struje i njihove komponente u električnom kolu posle povezivanja kondenzatora. Na osnovu prethodne analize zaključujemo da će struja kondenzatora uticati na smanjenje reaktivne komponente struje u električnoj mreži, odnosno reaktivna komponenta struje će se sa početne vrednosti $\underline{I}_{r1}$ smanjiti na vrednost $\underline{I}_{r2}$, dok će aktivna komponenta struje ostati nepromenjena jer se aktivno opterećenje u električnom kolu nije menjalo.

IV. Standardna merna metoda

Izvođenje laboratorijske vežbe iz popravke faktora snage sprovodi se neposredno u laboratoriji primenom ampermetra, voltmetra i vatmetra [6]. Principska šema za izvođenje ogleda prikazana je na Sl. 6.



Sl. 6. Električna šema za izvođenje ogleda primenom standardnih analognih mernih instrumenata.

Izvođenje laboratorijske vežbe sastoji se iz dva dela. Za prvi i drugi deo vežbe podešena vrednost napona ostaje nepromenjena. Prvi deo vežbe predstavlja merenje napona, struje i snage jednofaznog motora bez piključenog kondenzatora. Na osnovu izmerenih veličina određuje se faktor snage motora i fazni pomeraj između fazora napona i struje primenom izraza (2) i (3). Izmerene i izračunate veličine pregledno se beleže u tabelu.

$$cos\,\varphi_1 = \frac{P_{W1}}{U_1 \cdot I_1} \tag{2}$$

$$\varphi_1 = arccos(cos\,\varphi_1) \tag{3}$$

Drugi deo vežbe podrazumeva zatvaranje preklopnika P sa Sl. 6 i ponavljanje postupka merenja napona, struje, snage i struje kondenzatora. Izmerene vrednosti se pregledno beleže u tabelu i sprovode se izračunavanja nove vrednosti faktora snage i faznog pomeraja između fazora napona i struje primenom izraza (4) i (5).

$$cos\,\varphi_2 = \frac{P_{W2}}{U_2 \cdot I_2} \tag{4}$$

$$\varphi_2 = arccos(cos\,\varphi_2) \tag{5}$$

Nakon izvršenih merenja analiziraju se dobijeni rezultati. Porede se vrednosti faktora snage pre i posle povezivanja kondenzatora. Zatim se konstruiše (grafičkom metodom ili primenom odgovarajućih softverskih alata) vektorski dijagram napona i struja pre i posle povezivanja kondenzatora. Da bi se vektorski dijagram struja mogao pravilno nacrtati neophodno je sve struje nacrtati u istoj razmeri.

Neka od ograničenja standardne merne metode su sledeća: zahteva neposredno prisustvo u laboratoriji, samo neposredno merenje napona, struje i snage je duže zbog očitavanja analognih mernih instrumenata, ograničenja u primeni online-laboratorije. Zato je neophono primeniti savremene merne sisteme i uređaje koji će omogućiti izvođenje laboratorijske vežbe u online uslovima.

V. Primena instrumenta VMP 20 i rezultati merenja

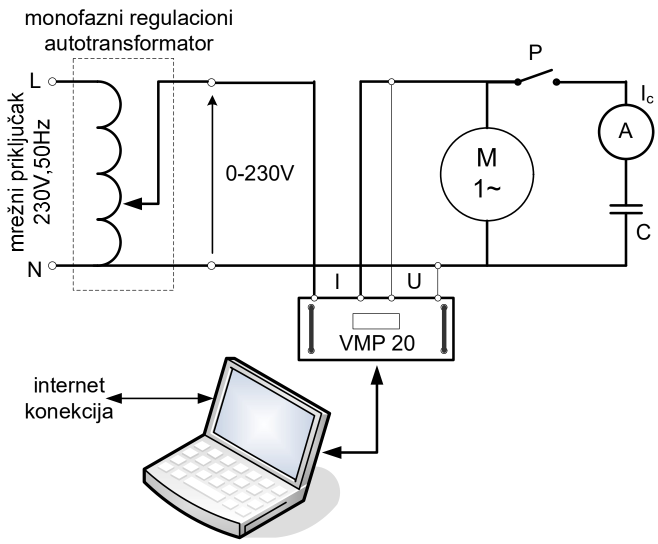
Laboratorijska vežba je tako koncipirana da je jasno definisan zadatak, data je električna šema i spisak ogledne opreme i kratko uputstvo za realizaciju laboratorijske vežbe. Učenik ili student, pre dolaska u laboratoriju, potrebno je da prouči tekst postavke laboratorijske vežbe i pripremiti se za nju.

U slučaju da se vežba radi u realnom vremenu, odnosno, onlajn, učenik ili student treba da bude pored svog računara u zakazanoj satnici. Neophodan preduslov je da postoji pristup internetu u laboratoriji za električna merenja.

Za rad u realnom vremenu preko interneta nastavnik ili saradnik su moderatori i definišu dinamiku izvođenja laboratorijske vežbe, dok učenik ili student preko video linka prate tok izvođenja laboratorijske vežbe.

U toku neposrednog merenja prozor softvera VMPCalc dostupan je putem video linka studentima ili učenicima i samostalno mogu izvršiti očitavanje merenih veličina.

Principska električna šema izvođenja ogleda uz primenu instrumenta VMP20 prikazana je na Sl. 7.

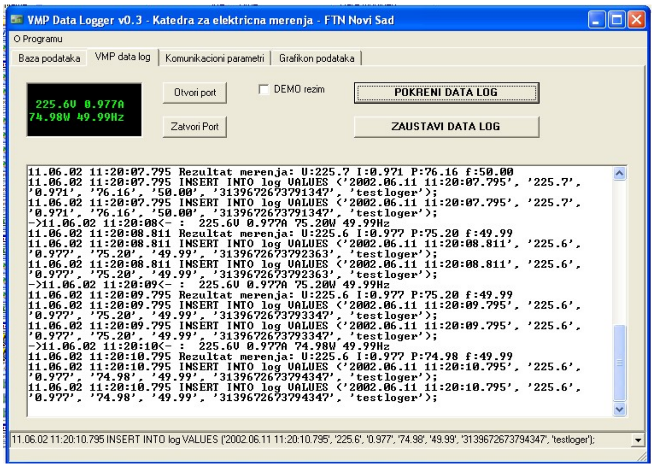


Sl. 7. Principski izgled električne šeme za izvođenje ogleda kompenzacije reaktivne snage primenom instrumenta VMP20.

Na Sl. 8. prikazan je izgled ispitne stanice u laboratoriji sa povezanom mernom opremom.

Sl. 8. Izgled ispitne stanice u laboratoriji sa povezanom mernom opremom i odgovarajućim računarom.

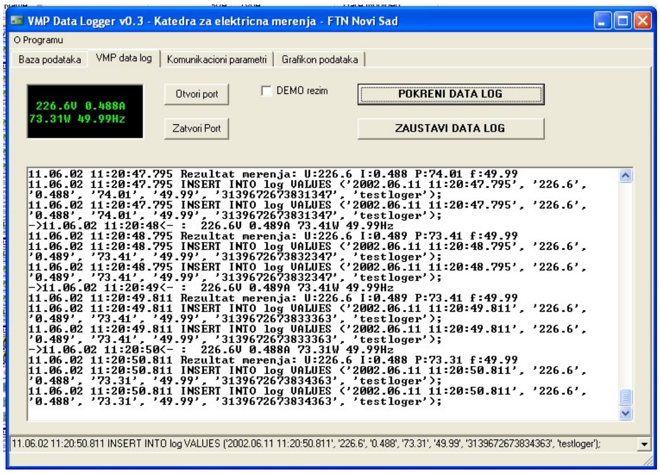
Tok izvođenja laboratorijske vežbe ogleda se u tome da se instrumentom VMP 20 meri napon, struja i snaga pre povezivanja kondenzatora. Primenom softvera VMPCalc na ekranu računara mogu se pratiti vrednosti merenih veličina što je prikazano na Sl. 9.



Sl. 9. Izgled prozora softvera VMPCalc pre povezivanja kondenzatora.

Zatim se preklopnik P sa Sl. 7. zatvori i izvrši se očitavanje pokazivanja instrumenta posle povezivanja kondenzatora, što je prikazano na Sl. 10.

Na osnovu rezultata merenja koji su prikazani na Sl. 9. i 10. uočavamo da se struja izvora napajanja smanjila sa vrednosti I_1=0.977 [A] na vrednost I_2=0.488 [A], dok je aktivna snaga koja se uzima iz električne mreže ostala približno jednaka jer se mehaničko opterećenje motora nije menjalo.



Sl. 10. Izgled prozora softvera VMPCalc posle povezivanja kondenzatora.

Smanjenje struje izvora upravo je posledica smanjenja reaktivne komponente struje električne mreže koju smo postigli povezivanjem kondenzatora paraleno električnom motoru.

S obzirom da se napajanje aparature vrši iz naizmeničnog izvora sa regulacionim autotransformatorom napajanje aparature možemo izvršiti i za nestandardne vrednosti napona električne mreže. U Tabeli I dati su rezultati merenja za nestandardnu vrednost napona izvora od 200 [V] .

TABELA I

TABELARNI PRIKAZ REZULTATA MERENJA ZA NAPON IZVORA 200 [V]

Pre povezivanja kondenzatora C		
U_1	200.4	[V]
I_1	0.714	[A]
P_1	52.86	[W]
$\cos\varphi_1$	0.36	
φ_1	68.9	[°]
I_C	0	[A]
Posle povezivanja kondenzatora C		
U_2	200.6	[V]
I_2	0.312	[A]
P_2	53.7	[W]
$\cos\varphi_2$	0.83	
φ_2	33.9	[°]
I_C	0.52	[A]
C	9	[μF]

Prilikom izvođenja drugog dela vežbe, paralelno elektromotoru povezana su dva međusobno paralelno povezana kondenzatora od 4.5 [μF].

Ovde je važno naglasiti da drugi deo vežbe može imati i drugačiji tok. Naime, moguće je zadati željenu vrednost faktora snage a zatim ostaviti učeniku ili studentu da sam proračuna potrebnu vrednost kapaciteta kondenzatora C_p, primenom izraza (6).

$$C_p = \frac{P_1}{\omega U_1^2}(tg\varphi_1 - tg\varphi_z) \qquad (6)$$

Pri čemu je značenje veličina u (6) sledeće: C_p-potrebna vrednost kapaciteta kondenzatora, ω-kružna učestanost električne mreže koja se jednostavno može odrediti jer instrument VMP20 u osnovnom prozoru meri i frekvenciju električne mreže, U_1-efektivna vrednost napona električne mreže, φ_1- fazni pomeraj između fazora napona i struje pre povezivanja kondenzatora, φ_z- zadata vrednost faznog pomeraja određena na osnovu zadate vrednosti faktora snage.

Kada odredimo potrebnu vrednost kapaciteta, iz grupe kondenzatora koji su na raspolaganju potrebno je izabrati odgovarajući, povezati ga u električno kolo, izvršiti drugo merenje, proveriti novu vrednost faktora snage i da li odgovara zadatoj vrednosti.

Završni deo vežbe podrazumeva diskusiju dobijenih rezultata, konstrukciju vektorskih dijagrama napona i struja elektromotora pre i posle povezivanja kondenzatora primenom geometrijske metode ili odgovarajućih softverskih alata (CAD softveri, GeoGebra i sl.) i sačinjavanje izveštaja.

VI. Zaključak

Cilj ovog rada je da približi jedan način izvođenja laboratorijske vežbe za popravku faktora snage (kompenzacija reaktivne snage). Prilikom realizacije laboratorijske vežbe korišćen je merni instrument VMP20 i odgovarajući softver VMPCalc. Na osnovu iskustva koje smo stekli u radu sa instrumentom možemo zaključiti da je izvođenje laboratorijke vežbe veoma elegantno, merenje se brzo vrši, softversko okruženje pruža organizaciju laboratorijske vežbe online, što omogućava učeniku ili studentu da prati tok izvođenja vežbe od kuće. Komunikacija preko video linka i brzo merenje instrumentom omogućavaju da se u terminu laboratorijskih vežbi ostavi sasvim dovoljno vremena za diskusiju i analizu dobijenih rezultata i da se uspostavi korelacija između teorijskih principa i praktičnih rezulatata kao i da se uspostavi njihova dodatna siteza.

Zahvalnica

Na ovom mestu želimo da se zahvalimo Vladimiru Vujičiću profesoru Fakulteta tehničkih nauka u penziji na velikoj podršci i savetima o mogućnostima primene instrumenta u neposrednoj laboratorijskoj praksi kao i pravcima primene u razvoju laboratorijskih merenja na daljinu i onlajn organizaciji laboratorijskih vežbi.

Literatura

[1] A. Juhas, "Hardverska i softverska podrška trofaznim merenjima snage i energije u ED mreži", master rad, Katedra za električna merenja, Fakultet tehničkih nauka u Novom Sadu, Novi Sad, Srbija, 2017.
[2] V. Vujičić, I. Župunski, Z. Mitrović, M. Sokola, "Merenje u tački i intervalu", Proc. of IMECO XIX, World Congress, Lisabon, Portugal, , pp. 1128-1132, No. 480, septembar, 2009.
[3] V. Vujičić, S. Milovančev, "Digitalni instrument za merenje srednje vrednosti proizvoda dava analogna periodična signala", YU patent, broj 48195, 1995.
[4] V. Pjevalica, V. Vujičić, "Further Generation of Low-Frequency True-RMS Instrument", IMTC, pp. 1008-1011, May, 2005.
[5] Branko Popović, "Osnovi elektrotehnike 2", Građevinska knjiga, Beograd, 1990.
[6] P. Dobanovački, S. Skoko, "Električna merenja-praktikum laboratorijskih vežbi", Elektrotehnička škola "Mihajlo Pupin", Novi Sad, Srbija, 2016.

Abstract

This paper presents a way of applying the measuring instrument VMP 20 in performing laboratory exercises in the subject of Electrical Measurements. The topic of the paper is reactive power compensation in a single-phase power supply system. The paper covers the basic theoretical principles of reactive power compensation and gives an overview of the laboratory exercise. The application of the VMP 20 instrument and the corresponding VMPCalc software has a special advantage in the organization of remote laboratory exercises - online, which is specifically indicated in the paper.

Application of the measuring instrument VMP 20 for a laboratory exercise - power factor correction

Isidora Sabadoš, Nemanja Vidović, Atila Juhas i Saša Skoko

Algoritam generisanja dvobitnih diterovanih Furijeovih bazisnih funkcija

J. Đorđević Kozarov, A. Juhas, P. Sovilj, *Member, IEEE*, V. Vujičić

***Apstrakt* — U radu je definisan algoritam generisanja dvobitnih diterovanih Furijeovih bazisnih funkcija (DDFBF) koje se koriste u SDDFT procesoru. Teorijski i eksperimentalno je potvrdjena njihova ortonormiranost sto je prikazano u radu.**

***Ključne reči* — Furijeova transformacija; Stohastika; Diter; Stohastička digitalna merna metoda.**

I. Uvod

ORTOGONALNE transformacije imaju veliki značaj u merenjima i obradi signala, a najstarija i najviše korišćena je Furijeova. Posebno je značajna brza Furijeova transformacija (FFT) zbog velike redukcije broja neophodnih množenja (najkritičnije aritmetičke operacije) i memorije za njeno izvršenje.

Napredak tehnologije računarskih komponenti, posebno što se tiče brzine procesora i kapaciteta memorije, postepeno ističu u prvi plan i DFT kao alternativu koja ima nekoliko prednosti u odnosu na FFT:

- Koeficijenti Furijeovog razvoja se izračunavaju međusobno nezavisno i, po potrebi, pojedinačno;
- Izračunavanje može da se vrši istovremeno sa odmeravanjem – kad se završi sekvenca i DFT može da bude završena;
- Jednostavna je paralelizacija u izračunavanju DFT.

Nijednu od gore nabrojanih osobina nema FFT.

Najnoviji razvoj u primeni DFT u merenjima snage i energije u elektrodistributivnoj (ED) mreži je doveo do realizacije namenskog stohastičkog digitalnog DFT (SDDFT) procesora čija je optimalna verzija – dvobitni SDDFT procesor [1].

Ključna operacija, MAK (Multiply and Accumulate), se sada izvršava jednostavno na jednostavnom hardveru – množač se sastoji od 4 dvoulazna „i“ kola i dva dvoulazna „ili“ kola, dok je akumulator – običan „up/down“ brojač. Jednostavna struktura hardvera za izvršavanje dvobitne MAK operacije, posebno ako se primeni u FPGA realizaciji SDDFT procesora, omogućuje visoku paralelizaciju izračunavanja DFT i time postizanja velikih brzina obrade analiziranog signala. Sa druge strane, i dvobitna stohastička digitalna A/D konverzija (SAADK) je izuzetno jednostavna, tako da su i paralelna merenja jednostavna [2].

U Literaturi [1] je dokazano da je u slučaju primene dvobitne SAADK optimalno primeniti dvobitne diterovane Furijeove bazisne funkcije. Problem generisanja dvobitnih diterovanih Furijeovih bazisnih funkcija (DDFBF) nije detaljnije elaboriran. Kod SDDFT visokog reda (preko 1000) taj problem postaje kritičan i zahteva posebnu pažnju.

II. Algoritam generisanja dvobitnih bazisnih funkcija (AGDBF)

1. U memoriji su odmerci Furijeovih bazisnih funkcija (FBF) u floating-point formatu (FP). Treba ih pretvoriti u celobrojni format (CF). Biramo broj bita celobrojnog formata – recimo 32 bita.
2. Pošto je jedan bit – bit znaka – izabrani odmerak FBF u FP formatu množimo sa 2^{31}, primenjujemo funkciju ROUND i smeštamo odmerak u LongInt (LI) formatu. Time smo spremili odmerak FBF za diterovanje.
3. Odgovarajući odmerak uniformnog šuma (UŠ) onda ima 31 bit, a dobijamo ga iz funkcije Random u FP formatu u opsegu (0,1),
4. Od odmerka UŠ oduzmemo 0,5 u FP formatu, pomnožimo sa 2^{30} i primenimo funkciju ROUND.
5. Celobrojno saberemo ta dva celobrojna odmerka. Rezultat ima 31 bit + bit znaka.
6. Detektujemo bit znaka u zbiru i primenimo na rezultat funkciju ABS.
7. Ako je najznačajniji bit apsolutne vrednosti zbira 1, apsolutna vrednost odgovarajuće vrednosti dvobitne FBF je 1, a ako nije – onda je apsolutna vrednost dvobitne FBF jednaka nuli (0).
8. Na apsolutnu vrednost dvobitne FBF se vrati bit znaka i *imamo kompletan dvobitni odmerak diterovane FBF*.
9. Kompletan dvobitni digitalni odmerak FBF (DDFBF) dalje koristimo u Furijerovom razvoju.

III. Eksperimentalna potvrda ortonormiranosti DDFBF

U radu [3] je strogo dokazana relacija (1):

$$\sigma_{\mathrm{e}}^2(2) = \frac{1}{N} \cdot \frac{(2g)^2}{T} \int_0^T |f_1(t) \cdot f_2(t)|\, dt -$$

Jelena Đorđević Kozarov – Elektronski fakultet, Univerzitet u Nišu, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: kozarov@elfak.ni.ac.rs).

Atila Juhas – Fakultet tehničkih nauka, Univerzitet u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: atila@ uns.ac.rs).

Platon Sovilj – Fakultet tehničkih nauka, Univerzitet u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: platon@ uns.ac.rs).

Vladimir Vujičić – Fakultet tehničkih nauka, Univerzitet u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija (e-mail: vujicicv@ uns.ac.rs).

$$-\frac{1}{N}\cdot\frac{1}{T}\int_0^T f_1^2(t)\cdot f_2^2(t)\,dt \qquad (1)$$

$\sigma_{\bar{e}}$ je preciznost merenja srednje vrednosti proizvoda dve funkcije, $y_1 = f_1(t)$ i $y_2 = f_2(t)$, dvobitnom SDMM na vremenskom intervalu $[0, T]$. U ED mreži, T je perioda mrežnog napona.

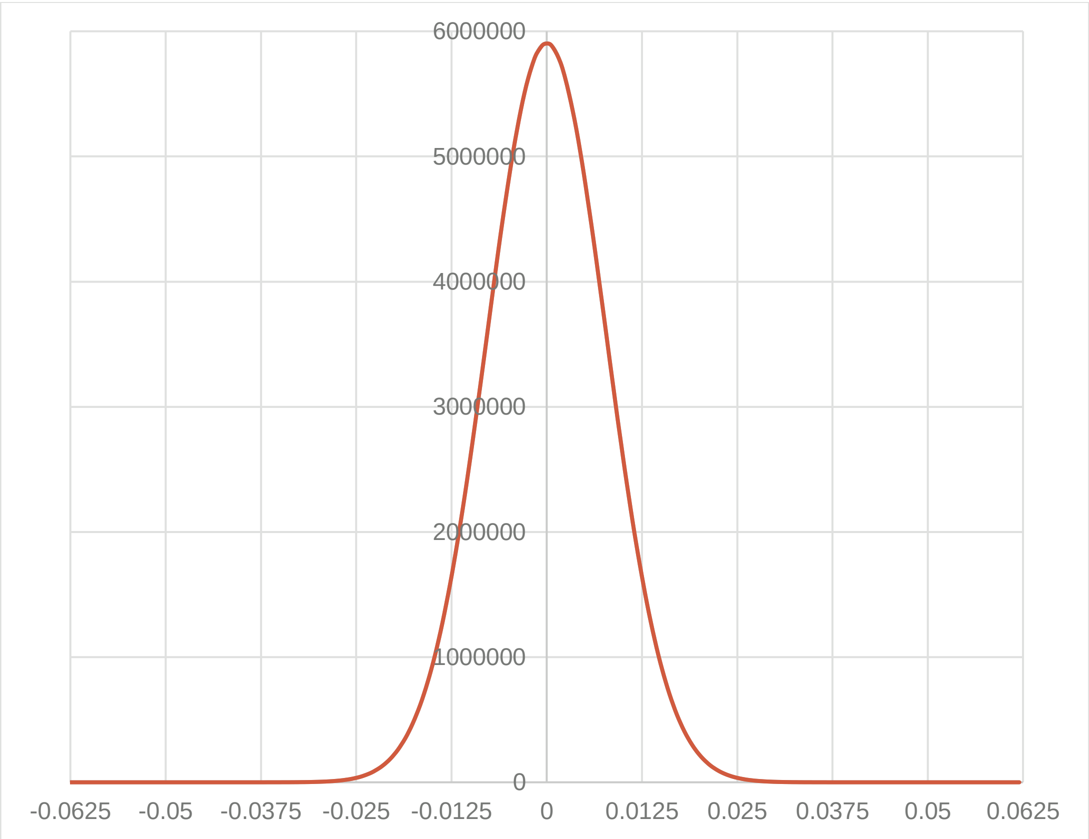
Normiranost DDFBF funkcija je eksperimentalno potvrđena, što je prikazano na slici 1. Teorijsko σ na toj slici je, zapravo, $\sigma = \sigma_{\bar{e}}(2)$ za *g=1*, $f_1(t) = f_2(t) = \sin(n\omega t)$ ili $\cos(n\omega t)$, $\{n = 1, 2, \ldots, 32\}$, na 50 Hz sa 2048 semplova po periodi, tj. N = 2048. $\sigma_{\bar{e}}$ je usrednjeno σ na kompletnom ansamblu od 64 norme u 236800000 tačaka na vremenskoj osi. $\bar{e}$ je srednja vrednost greške svakog proizvoda odmeraka u 236800000 tačaka.

Ortogonalnost DDFBF funkcija je eksperimentalno potvrđena, što je prikazano na slici 2. Teorijsko σ na toj slici je, zapravo, $\sigma = \sigma_{\bar{e}}(2)$ za *g=1*, $f_1(t) = \sin(n\omega t)$, $f_2(t) = \sin(m\omega t)$ ili $f_1(t) = \cos(n\omega t)$, $f_2(t) = \cos(m\omega t)$, $n, m \in \{1, 2, \ldots, 32\}$ i $n \neq m$ na 50 Hz, sa 2048 semplova po periodi, tj. N = 2048. $\sigma_{\bar{e}}$ je usrednjeno σ na kompletnom ansamblu od $2\cdot\left[\binom{32}{2} - 32\right] = 1742$ ortogonalne kombinacije u 236800000 tačaka na vremenskoj osi. $\bar{e}$ je srednja vrednost grešaka svakog proizvoda odmeraka u 236800000 tačaka.

Ortogonalnost DDFBF funkcija je eksperimentalno potvrđena, što je prikazano na slici 3. $\sigma = \sigma_{\bar{e}}(2)$ za *g=1*, $f_1(t) = \sin(n\omega t)$, $f_2(t) = \cos(m\omega t)$, $n, m \in \{1, 2, \ldots, 32\}$ i $n \neq m$, na 50 Hz, sa 2048 semplova po periodi, tj. N = 2048. Ovde, $\sigma_{\bar{e}}$ je usrednjeno σ na kompletnom ansamblu od $64^2 - 2\cdot\left[\binom{32}{2} - 32\right] - 64 = 2290$ ortogonalnih kombinacija u 236800000 tačaka na vremenskoj osi, dok je $\bar{e}$ srednja vrednost grešaka svakog proizvoda odmeraka u 236800000 tačaka.

TABELA I
NUMERIČKA POTVRDA NORMIRANOSTI DDFBF

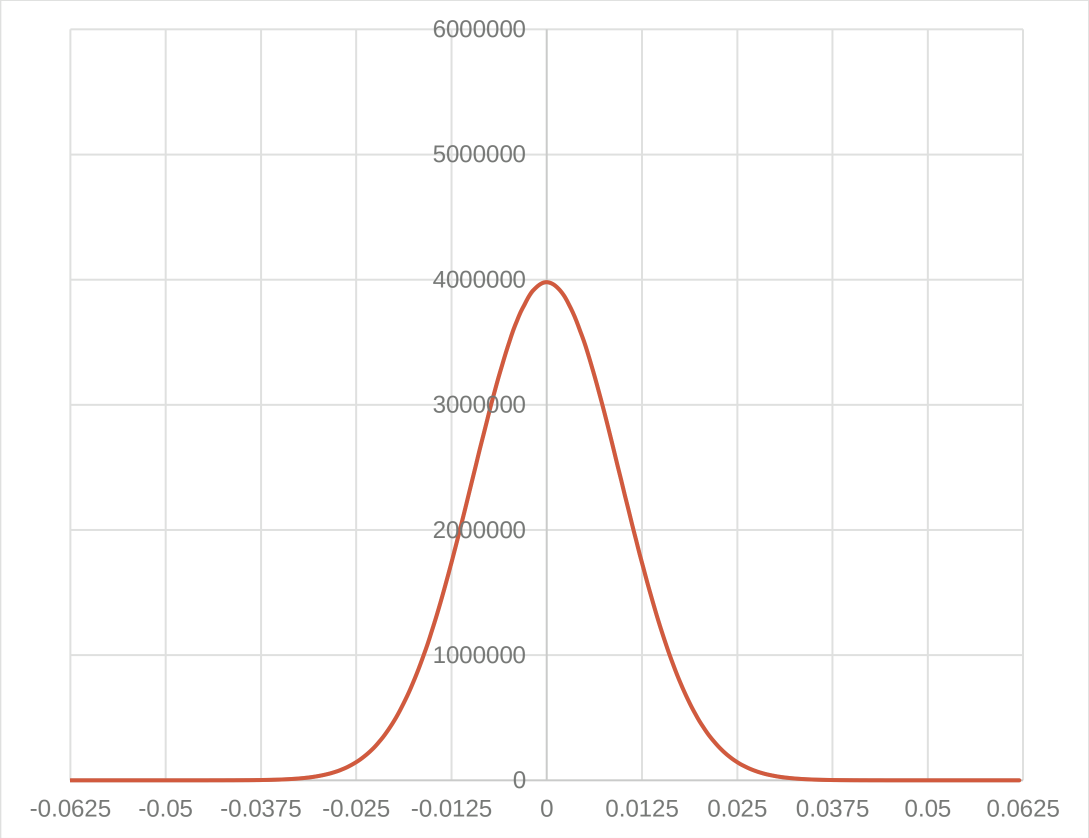
Broj semplova	Slučaj	Teorijsko **σ** formula	Teorijska vrednost **σ**	Eksperimentom dobijeno **σ**	$\frac{\sigma}{\sigma_{\bar{e}}}$	$\bar{e}$	$\frac{\bar{e}}{\sigma_{\bar{e}}}$	Broj tačaka
2048	a)	$\sigma_{\bar{e}} = \frac{1}{4}\sqrt{\frac{2}{N}} = \sqrt{\frac{1}{8N}}$	0.0078125	0.007814339	1.000235375	-5.37027E-07	-6.87394E-05	236800000



Sl.1. Grafička potvrda normiranosti DDFBF

TABELA II
NUMERIČKA POTVRDA ORTOGONALNOSTI DDFBF - 1

Broj semplova	Slučaj	Teorijsko **σ** formula	Teorijska vrednost **σ**	Eksperimentom dobijeno **σ**	$\frac{\sigma}{\sigma_{\bar{e}}}$	$\bar{e}$	$\frac{\bar{e}}{\sigma_{\bar{e}}}$	Broj tačaka
2048	b)	$\sigma_{\bar{e}} = \sqrt{\frac{8}{\pi} - 1} \cdot \sqrt{\frac{1}{8N}}$	0.009715431	0.009713238	0.999774324	-6.28933E-07	19840000005	236800000



Sl. 2. Grafička potvrda otogonalnosti DDFBF - 1.

.

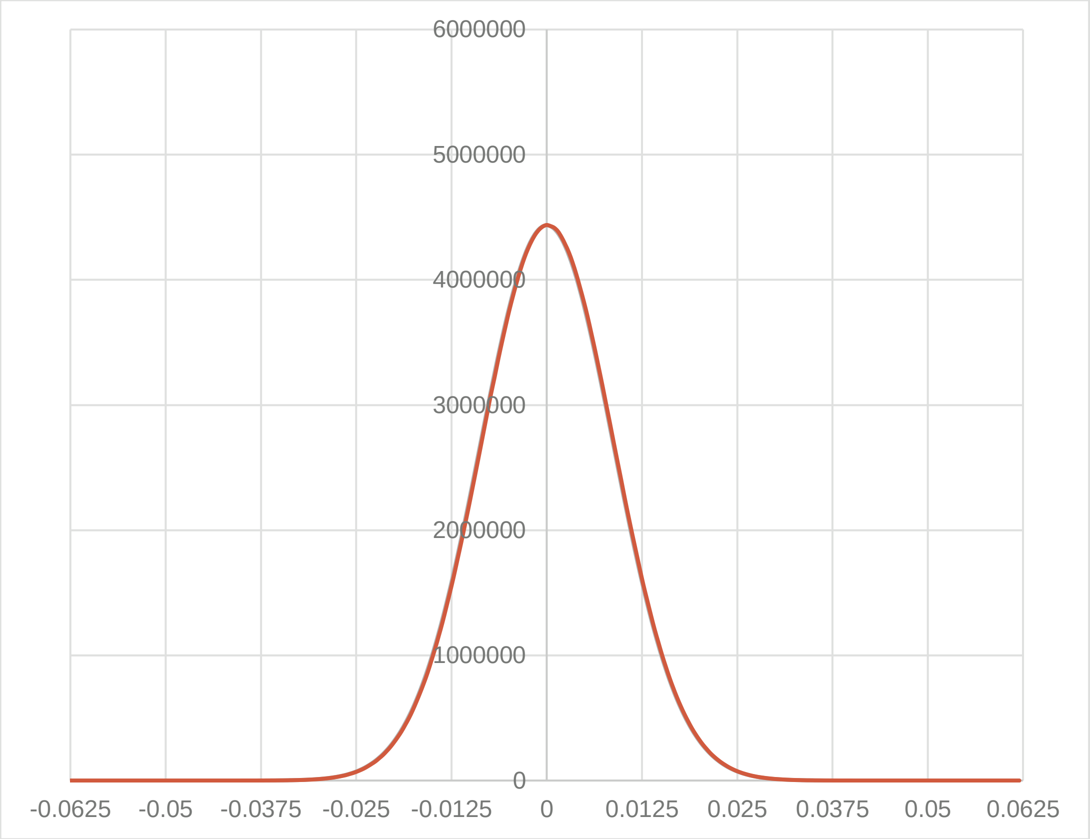
IV. ZAKLJUČAK

U radu je definisan algoritam generisanja digitalnih dvobitnih Furijeovih bazisnih funkcija (DDFBF). Na osnovu teorijskog kriterijuma granične preciznosti norme i ortogonalnosti, eksperimentalno je potvrđena ortonormiranost DFT sa 32 harmonika. Time je teorijski i eksperimentalno dokazana korektnost opisanog algoritma generisanja. Eksperiment je bio vrlo detaljan i sveobuhvatan i za normiranost i za oba tipa ortogonalnosti. Izveden je u po 236800000 tačaka u svakoj od tri varijante ortonormiranosti. Slaganje teorijske i eksperimentalne preciznosti je vrlo prihvatljivo i sa velikom pouzdanošću stoji tvrdnja da je predloženi algoritam generisanja DDFBF – korektan.

.

TABELA III
NUMERIČKA POTVRDA ORTOGONALNOSTI DDFBF – 2

Broj semplova	Slučaj	Teorijsko **σ** formula	Teorijska vrednost **σ**	Eksperimentom dobijeno **σ**	$\frac{\sigma}{\sigma_{\bar{e}}}$	$\bar{e}$	$\frac{\bar{e}}{\sigma_{\bar{e}}}$	Broj tačaka
2048	c)	$\sigma_{\bar{e}} \leq \sqrt{2}\sqrt{\frac{8}{\pi} - 1} \cdot \sqrt{\frac{1}{8N}}$	0.013739694	0.008708729	0.633837169	1.22603E-04	8.92326E-03	198400000



Sl.3. Grafička potvrda ortogonalnosti DDFBF – 2

.

ZAHVALNICA

Ovaj rad je podržan od strane Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije.

LITERATURA

[1] D. Pejic et al., "Stochastic Digital DFT Processor and Its Application to Measurement of Reactive Power and Energy", Measurement, vol. 124, pp. 494-504, Aug. 2018.

[2] P Sovilj et al. "Synergy and Completeness of Simple A/D Conversion and Simple Signal Processing", J. Phys.: Conf. Ser., vol. 1379, paper no. 012064, July 2019.

[3] V. Vujicic et al., "Stochastic Measurement of Wind Power using a Two-bit A/D Converter", Measurement, vol. 152, paper no. 107184, Feb. 2020

ABSTRACT

The paper defines an algorithm for generating two - bit ditered Fourier basis functions (DDFBF) used in an SDDFT processor. Theoretically and experimentally their orthonormalization was confirmed, which is shown in the paper.

Algorithm for two-bit ditered Fourier basis functions generating

J. Đorđević Kozarov, A. Juhas, P. Sovilj, V. Vujičić

Optimalna rezolucija stohastičkih embedid sistema

Dragan Pejić, Aleksandar Radonjić, Vladimir Vujičić

***Sažetak*-U radu se elaborira primena stohastičke digitalne merne metode (SDMM) u embedid sistemima. U literaturi je dokazano da je optimalna rezolucija SDMM trobitna i da je optimalni brojni sistem za primenu u obradi takođe trobitni. Ove činjenice su motivisale autore da analiziraju optimalne rezolucije stohastičke A/D konverzije, stohastičke obrade i stohastičke D/A konverzije.**

***Ključne reči*-Stohastička merenja, embedid sistemi, analogno -digitalna konverzija, digitalno-analogna konverzija.**

I. UVOD

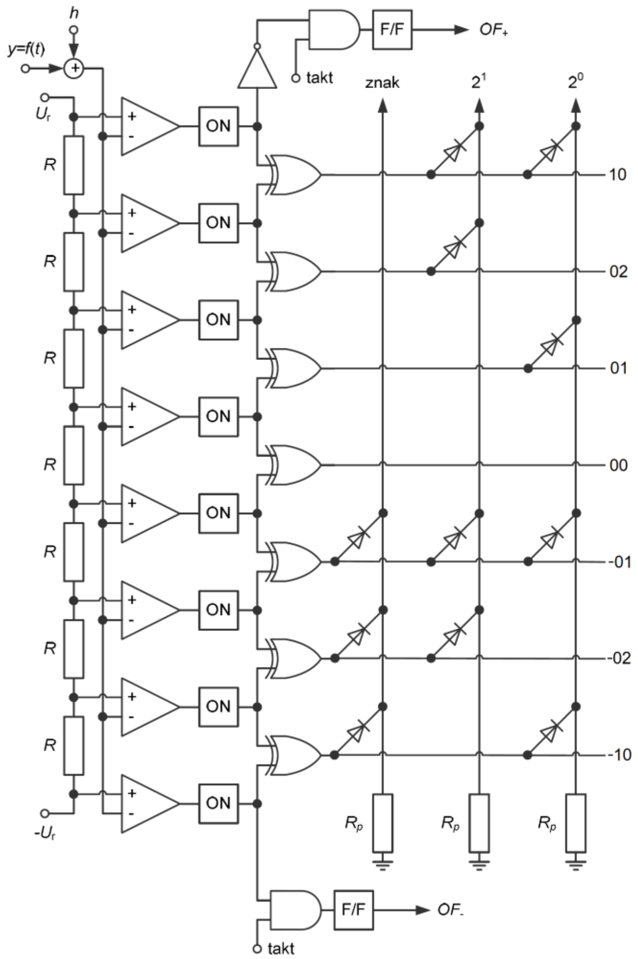
U zadnjih nekoliko godina, u akademskim krugovima, ali i u industriji, veliku pažnju privlači koncept stohastičkog računanja (SR). Najsveobuhvatniji prikaz ove teme dat je u [1]. Autori tog rada su prikazali načine primene stohastičkog računanja (SR) u embedid sistemima, ali i na koji način se celobrojna aritmetika može kombinovati sa stohastičkom analogno-digitalnom (A/D) konverzijom [2]. Cilj ovog rada jeste da se pomenute teme prodube, odnosno da se da dodatni doprinos teoriji stohastičkih embedid sistema.

II. STOHASTIČKA ANALOGNO-DIGITALNA KONVERZIJA

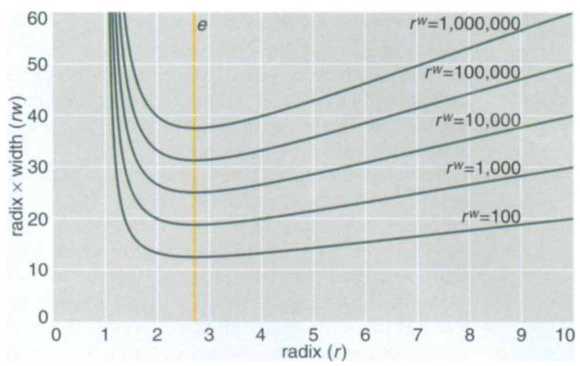
Kao što je već navedeno, u [1] je dat detaljan pregled SR metoda. Međutim, zbog svoje sveobuhvatnosti, pomenuti rad obuhvata teme koje se, u suštini, ne bave SR, već nekim drugim temama, kao što je stohastička A/D konverzija. Primer za to jeste i rad [2], gde se analogni uniformni diter dodaje ulaznom signalu pre njegove digitalizacije pomoću dvobitnog A/D konvertora (slika 1). Dakle, stohastika je prisutna na ulazu u sistem, dok su procesi obrade signala potpuno deterministički. Rezultat ovakvog pristupa jeste činjenica da ulazna veličina (signal) ima mali nivo slučajnosti, ali se zato meri sa velikom preciznošću i tačnošću.

Stohastička A/D konverzija je još uočljivija u [3]-[6], gde se koriste višebitni A/D konvertori. U svim slučajevima, blokovi za obradu signala su deterministički, iako bi se mogli realizovati i na stohastički način. Kada bi se tako nešto uradilo, pojednostavila bi se hardverska struktura uređaja, ali bi se značajno smanjila preciznost merenja zbog dužeg vremena računanja. S obzirom da se stohastika primenjuje na ulaznim podacima, korišćenje aritmetike sa pokretnim zarezom ne bi imalo smisla, jer se adekvatna preciznost rezultata može postići samo u dovoljno dugom vremenskom periodu. To se, inače, već postiže celobrojnom aritmetikom, što su potvrdili rezulatati simulacija [3]-[6].

Dragan Pejić, Fakultet tehničkih nauka Novi Sad, Novi Sad, Srbija (i-mejl: pejicdra@uns.ac.rs)
Aleksandar Radonjić, Institut tehničkih nauka SANU, Beograd, Srbija (i-mejl: sasa_radonjic@yahoo.com)
Vladimir Vujičić, Fakultet tehničkih nauka Novi Sad, Novi Sad, Srbija (i-mejl: vujicicv@uns.ac.rs)

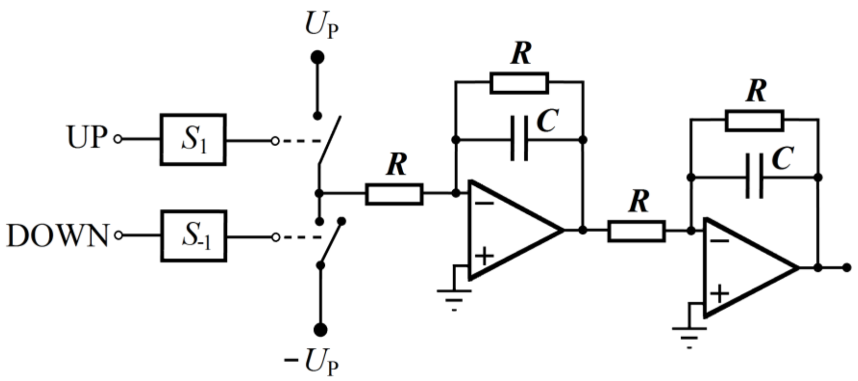


Sl.1. Šema trobitnog diterovanog A/D konvertora sa binarnim i ternarnim izlazom.



Sl. 2. Zavisnost cene primene od brojne osnove r i opsega računara r^w [8].

U [6] je pokazano kako kombinacija dvobitne stohastičke A/D konverzije i celobrojne aritmetike (visoke rezolucije) omogućava precizno merenje veoma važne vrednosti u električnoj mreži: osnovne frekvencije. Prednost pomenutog pristupa je uklanjanje celobrojnog množenja u FIR filtru. Ova činjenica nas navodi na zaključak da u embedid sistemima mogu postojati četiri načina digitalizacije i obrade signala: A/D konverzija i determinističko računanje (DR),

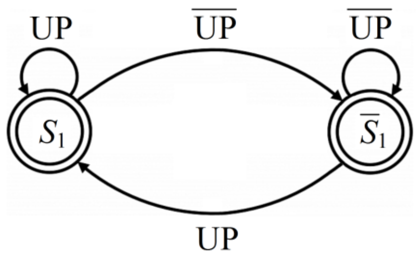


Sl. 3. Inherentno dvobitni stohastički analogni AC / DC transfer [7].

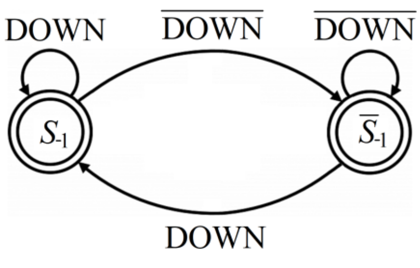
A/D konverzija i SR, stohastička A/D konverzija i DR, i stohastička A/D konverzija i SR. Poslednja kombinacija ima najjednostavniji hardver, i veoma je pogodna za primenu u embedid sistemima koji mere, nadgledaju ili regulišu spore procese (hemijski i biološki procesi) u poljoprivredi i zaštiti životne sredine. Jednostavan hardver znači i da minimalno korišćenje energije, što embedid sisteme, zasnovane na SR i stohastičkoj A/D konverziji čini pogodnim za realizaciju u formi autonomnog baterijski-napajanog ASIC čipa.

III. Optimalna Stohastička Obrada

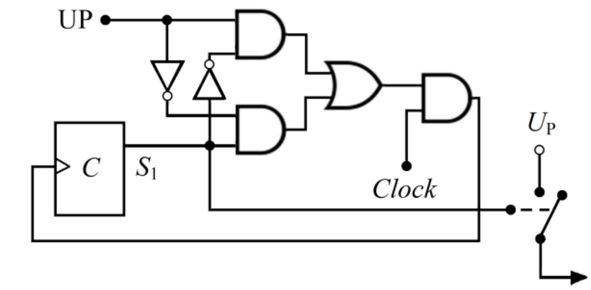
Još od rane faze razvoja računara povela sa diskusija o optimalnom brojnom sistemu koji bi se koristio u računarima. Rezime tih diskusija je dat u članku [8]. Eksplicitno je utvrđeno da je brojni sistem sa osnovom *e* (*e* ≈ 2,718) najekonomičniji sa stanovišta zapisa u memoriji, pretraživanja i obrade. Kako brojna osnova mora biti ceo broj, najbliži ceo broj je 3, pa svi zaključci koji se odnose na brojnu osnovu *e* se ostvaruju usvajanjem brojne osnove (modula, radix-a) 3. To je prikazano na slici 2. Vrlo rano je razvijena i ternarna Bulova algebra, tako da, generalno, ternarni brojni sistem i ternarna obrada su najekonomičniji sa stanovišta obrade.



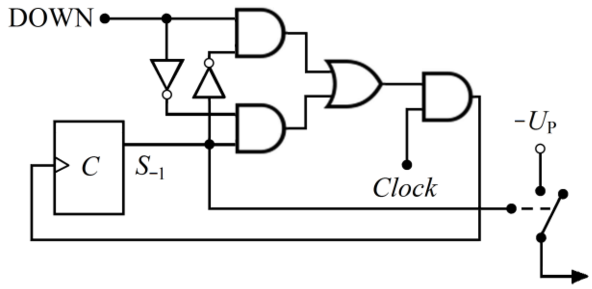
Sl. 4. Grafički prikaz mašine stanja S_1 [7].



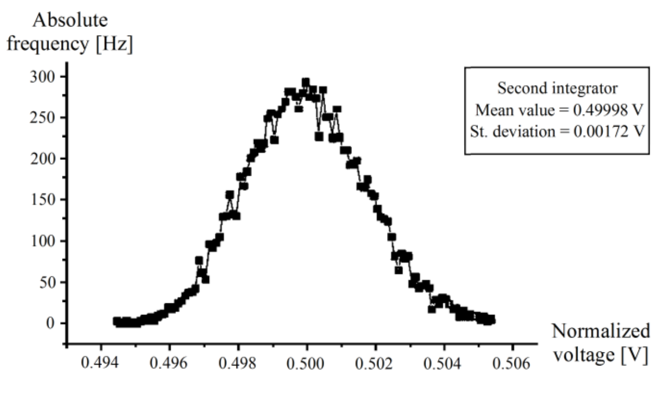
Sl. 5. Grafički prikaz mašine stanja S_{-1} [7].



Sl. 6. Logički prikaz mašine stanja S_1 [7].



Sl. 7. Logički prikaz mašine stanja S_{-1} [7].



Sl. 8. Izlaz drugog analognog integratora u slučaju sinusoidalnog ulaza u uređaj [7].

Da li je bolja stohastička obrada, u prvom redu računanje, binarna ili ternarna, je jednostavno utvrditi po analogiji sa njihovim determinističkim analogonima - bolja je ternarna.

Kako je stohastički embedid sistem vrlo složen tehnički entitet, koji se sastoji od tri podsistema: ulaza (A/D konverzija), bloka za obradu (realizacija algoritama i računanje) i izlaza (D/A konverzija), a na osnovu poznate Pontijaginove teoreme - optimum sistema se nikada ne poklapa sa optimumom bilo kog njegovog podsistema - gotovo je neverovatan rezultat istraživanja [9] da je ternarni izlaz iz SAADK sa slike 1 takođe optimalan. Već na osnovu te činjenice vredi razmisliti i o ternarnom SAADK. U tom slučaju je kompletan stohastičke embedid sistem ternaran. To je nepobitna činjenica, a najnoviji razvoj FPGA tehnologije nudi velike mogućnosti njegove realizacije.

IV. Stochastička Digitalno-Analogna Konverzija

Treća važna funkcija embedid sistema je digitalno-analogna (D/A) konverzija. U stohastičkim embedid sistemima ona se lako izvršava pomoću analognih filtera. To je sintetički prikazano na slici 3, dok su ostali detalji prikazani na slikama 4-8 u slučaju dvobitnog stohastičkog embedid sistema. U [7] je pokazano da je za postizanje jednostavnog, robusnog i preciznog AC/DC transfera dovoljno koristiti dvobitni stohastički A/D konvertor, dvobitni deterministički blok za obradu signala i jednostavan inherentno stohastički dvobitni D/A kovertor na ulazu u analogni filter drugog reda (redno vezana dva analogna integratora sa gubicima). Potpuno analogno se može uraditi ternarna stohastička DA konverzija trobitnim izlazima iz ternarnog bloka za obradu.

V. Zaključak

U radu je pokazano da stohastička ternarna A/D konverzija znatno pojednostavljuje hardver embedid sistema. Istraživanje i analiza metoda, zasnovanih na ternarnoj stohastičkoj A/D konverziji, dovelo bi do dodatnog unapređenja embedid sistema. Prednosti bi bile slične onima kod ternarnog SR-a. Sinergijski efekat kombinovane primene ternarne SR i ternarne stohastičke A/D konverzije nudi velike mogućnosti za dodatno pojednostavljenje hardvera embedid sistema sa svim prednostima koje ova činjenica podrazumeva. Ternarna stohastička DA konverzija je potpuno ostvarljiva na analogan način binarnoj.

Literatura

[1] A. Alaghi and J. P. Hayes, "Survey of Stochastic Computing," *ACM Trans. Embedded Comput. Syst.*, vol. 12, no. 2s, pp. 1-19, May 2013.

[2] D. Pejic and V. Vujicic, "Accuracy Limit of High Precision Stochastic Watt-Hour Meter," *IEEE Trans. Instrum. Meas.*, vol. 49, no. 3, pp. 617-620, Jun. 2000.

[3] V. Vujicic, "Generalized Low-Frequency Stochastic True RMS Instrument," *IEEE Trans. Instrum. Meas.*, vol. 50, no. 5, pp. 1089-1092, Oct. 2001.

[4] V. Pjevalica and V. Vujicic, "Further Generalization of the Low-Frequency True-RMS instrument," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 3, pp. 736-744, Mar. 2010.

[5] P. Sovilj, S. Milovancev, and V. Vujicic, "Digital Stochastic Measurement of a Nonstationary Signal with an Example of EEG Signal Measurement," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 9, pp. 3230–3232, Sep. 2010.

[6] A. Radonjic, P. Sovilj, and V. Vujicic, "Stochastic Measurement of Power Grid Frequency Using a Two-Bit A/D Converter," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 1, pp. 56–62, Jan. 2014.

[7] V. Vujicic and D. Pejic, "Inherently Digital Stochastic Analog AC to DC Transfer," *in: Proc. IMTC 2001*, vol. 3, pp. 2070-2073, May 2001.

[8] Brian Hayes, "Computing Science: Third Base,"*American Scientist*, vol. 89, no. 6, pp. 490-494, Nov.-Dec. 2001.

[9] M. Urekar, Contribution to the optimization of digital measurements performance (in Serbian), Ph.D. dissertation, Faculty of Technical Sceinces, University of Novi Sad.

ABSTRACT

The paper elaborates the application of stochastic digital measurement method (SDMM) in embedded systems. It has been proven in the literature that the optimal SDMM resolution is three-bit and that the optimal number system for application in signal processing is also three-bit. These facts motivated the authors to analyze the optimal resolutions of stochastic A/D conversion, stochastic processing, and stochastic D/A conversion.

Optimal resolution of stochastic embedid systems

Dragan Pejic, Aleksandar Radonjic and Vladimir Vujicic

Идејни пројекат генератора аналогног дискретног униформног шума

Бојан Вујичић, Драган Пејић, Александар Радоњић, Владимир Вујичић

***Сажетак*-У раду се предлаже идеја генератора аналогног дискретног униформног шума (ГАДУШ) за примену у стохастичкој дигиталној мерној методи. Предложеним решењем, за разлику од стандардног, које користи генератор случајних бројева и прецизни и тачни дигитално-аналогни конвертор (ДАК), се превазилази проблем ограничене резолуције ДАК-а и његов узак пропусни опсег.**

***Кључне речи*-Стохастичка дигитална мерна метода, генератор аналогног дискретног униформаног шума, дигитално-аналогна конверзија.**

I. Увод

Основни мотив за ово истраживање и писање овог рада је потреба за квалитетним генератором аналогног дискретног униформног шума (ГАДУШ) који би нашао примену у стохастичким дигиталним мерењима (СДМ) [1]-[6]. У [5] је дата строга дефиниција једноставне аналогно-дигиталне (А/Д) конверзије и једноставне обраде сигнала у оквиру једноставне СДМ методе (СДММ). Обе су или једнобитне или двобитне. Ова чињеница имплицира једноставан хардвер и свега пар извора систематске грешке у једноставној СДММ који се лако идентификују и елиминишу [3], [7]. Изузетна тачност је најважнија перформанса једноставне СДММ.

С друге стране, сваки нови бит у СДММ је еквивалентан примени приближно четири пута брже технологије [8]. Ово убрзање се плаћа дупло сложенијим хардвером, а самим тим и дупло већим бројем извора систематске грешке. Оптимална резолуција СДММ је три бита [8]. Тада ефективно убрзање технологије износи девет пута у односу на двобитну СДММ. Привлачна је, стога, идеја врло тачног мерења тробитном СДММ, или, уопште, вишебитном СДММ. Та идеја није предмет овог рада.

У [9] је дат кратак преглед инструмената у којима је примењена једноставна СДММ. Општа особина свих тих инструмената је велика тачност али не само она. Једноставан хардвер имплицира и велике могућности паралелног мерења и паралелне обраде сигнала. У [9] је наведена могућност детекције основног и најзначајнијег вишег хармоника у клизећем прозору од половине периоде мрежног напона, док је у [10] описана примена једноставне СДММ у решавању проблема редукције велике количине података у електро-дистрибутивној (ЕД) мрежи.

Бојан Вујичић, Факултет техничких наука Нови Сад, Нови Сад, Србија (и-мејл: bojanvuj@uns.ac.rs)

Владимир Вујичић, Факултет техничких наука Нови Сад, Нови Сад, Србија (и-мејл: vujicicv@uns.ac.rs)

Александар Радоњић, Институт техничких наука САНУ, Београд, Србија (и-мејл: sasa_radonjic@yahoo.com)

Драган Пејић, Факултет техничких наука Нови Сад, Нови Сад, Србија (и-мејл: pejicdra@uns.ac.rs)

У свим варијантама СДММ је неопходан елемент ГАДУШ. Поновимо још једном општу дефиницију СДММ у којој се то и експлицитно види. СДММ је: дигитално мерење средње вредности производа два сигнала на временском интервалу када се сваком од сигнала пре АД конверзије дода шум униформне расподеле у опсегу +/- половине кванта АД конвертора. Шумови на оба канала морају бити међусобно некорелисани а средња вредност сваког од њих мора бити нула. Усредњавају се, не дигитални одмерци, него производи парова одмерака. Све остало је или специјалан случај горе наведене дефиниције или њена директна генерализација.

II. Пројектна идеја ГАДУШ-а

Термички или Никвистов шум у отпорнику има Гаусову густину вероватноће расподеле:

$$f(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2} \tag{1}$$

Претпоставимо, сада, да је шум x могуће линеарно појачати κ пута у опсегу од интереса (бар у 3σ). Тада за појачани шум важи:

$$f(kx) = \frac{1}{k \cdot \sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{kx}{k\sigma}\right)^2} \tag{2}$$

У $x \le x_g \approx 0$ околини $x = 0$, $kx = 0$ је фактор одступања Q Гаусове расподеле од униформне:

$$Q = \frac{1}{2} \cdot \frac{f(k\cdot 0) - f(k\cdot x_g)}{f(k\cdot 0)} \approx \frac{1}{4} \cdot \frac{x_g^2}{\sigma^2} \tag{3}$$

У Табели 1 су приказане карактеристичне вредности x_g, односно kx_g у зависности од изабраног фактора одступања Q (напомена: $k = 10^5$ и $\sigma = 10^{-6} V$).

ТАБЕЛА 1
ЧЕТИРИ КАРАКТЕРИСТИЧНЕ ВРЕДНОСТИ x_g И kx_g

Q	x_g [V]	kx_g [V]
0.001	$6.32\cdot10^{-8}$	$6.32\cdot10^{-3}$
0.0001	$2\cdot10^{-8}$	$2\cdot10^{-3}$
0.00001	$6.32\cdot10^{-9}$	$6.32\cdot10^{-4}$
0.000001	$2\cdot10^{-9}$	$2\cdot10^{-4}$

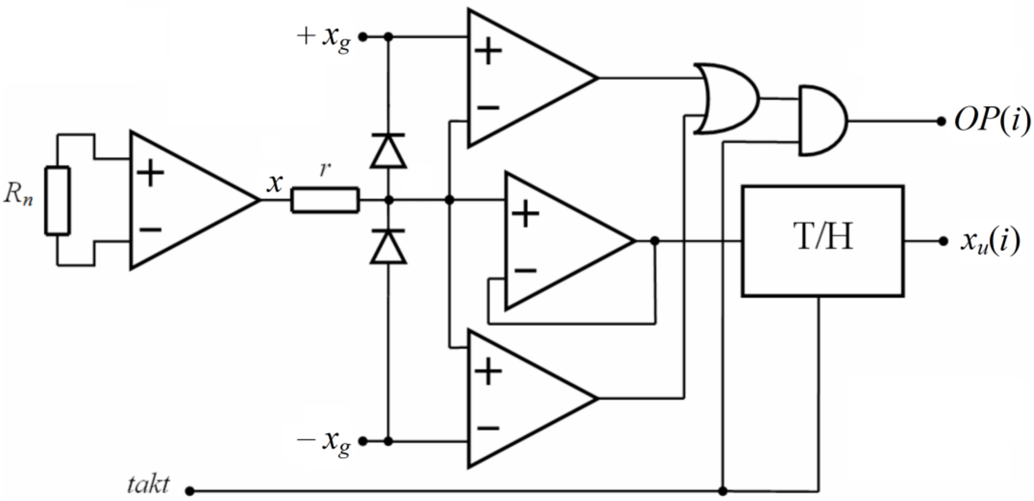
III. Поставка проблема

Теорија и досадашња искуства су показала да ГАДУШ мора имати следеће особине:

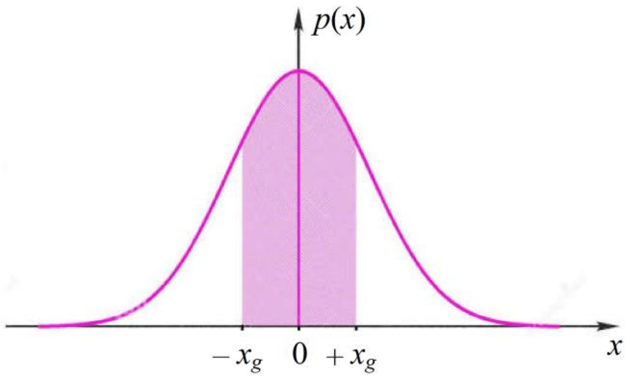
1) ГАДУШ мора да генерише напон стварно униформне расподеле,

2) Одмерци ГАДУШ-а морају бити међусобно некорелисани – аутокорелациона функција излаза ГАДУШ-а мора бити нула,

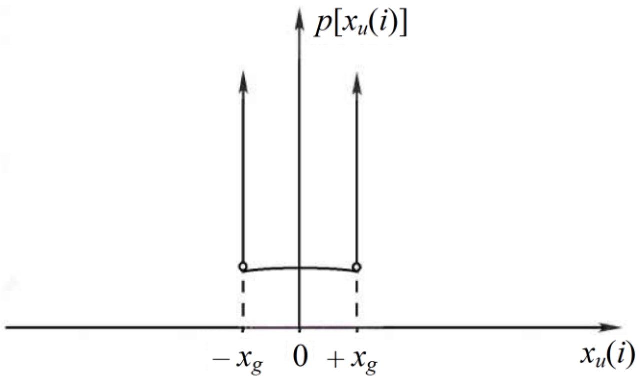
3) Одмерци ГАДУШ морају бити у опсегу једног кванта стохастичког адиционог А/Д конвертора



Слика 1. Предлог новог ГАДУШ-а.



Слика 2. Гаусова функција густине расподеле вероватноће (облик Никвистовог отпорничког шума).



Слика 3. Функција густине расподеле вероватноће на излазу из Т/Н кола.

(СААДК),

4) Одмерци морају бити распоређени униформно у интервалу времена,

5) Одмерци морају бити аналогни морају да имају бесконачно велику резолуцију).

Само такав ГАДУШ може заменити постојеће решење засновано на генератору случајних бројева и ДА конвертору и, евентуално, превазићи његова ограничења. У постојећем решењу је уско грло ДА конвертор, посебно, његов пропусни опсег [11].

Аналогни дискретни случајни одмерци на његовом излазу се свраставају у две класе: а) неуниформни случајни одмерци када излазни бинарни класификациони сигнал OP(i) има вредност 0 (лаж) и б) униформни случајни одмерци када излазни бинарни класификациони сигнал има вредност 1 (истина). За дигитално мерење и дигиталну обраду сигнала се користе само одмерци из класе б). Одмерци из класе а) се игноришу – не обрађују се.

Извор шума је отпорник а термички шум у отпорнику има нормалну (Гаусову) функцију густине расподеле вероватноће са средњом вредношћу нула. У довољно уским симетричним границама $|x| < |x_g|$ око нуле Гаусова расподела је униформна. Термички шум из отпорника се снажно напонски појачава и преко редног отпорника доводи на симетрични двострани диодни ограничивач. Када је тренутна вредност појачаног термичког шума унутар граница $|x| < |x_g|$, диоде су закочене и излазни бинарни класификациони сигнал има вредност 1. Тада тренутна вредност појачаног термичког шума иде у даљу обраду. Када је тренутна вредност појачаног термичког шума изван горенаведених граница, излазни бинарни класификациони сигнал има вредност 0, сама тренутна вредност појачаног термичког шума се преко редног отпорника своди на граничну вредност увећану за пад напона на диоди и игнорише се – не иде у даљу обраду.

Важан елемент у овој идеји је коло за праћење и задршку (енгл. Track & Hold – T/H). Ако су симетричне границе диодног ограничивача уске, тј. ако $|x_g| \to 0$, онда T/H коло може да ради на врло високој учестаности и тиме се превазилази проблем уског пропусног опсега ГАДУШ-а. Чињеница да су одмерци аналогни имплицира да је у овом ГАДУШ-у превазиђен и проблем ограничене резолуције претходних решења. Дигитални генератор такта дефинише посматране дискретне тренутке у обе класе одмерака.

IV. Дискусија

Са слика 2 и 3 је јасно да уколико су симетричне границе $|x_g| \to 0$ уже, расподела приказана на слици 3 је ближа униформаној у области $|x| < |x_g|$. Више технолошких параметара утиче на употребљивост ове идеје реализације ГАДУШ-а. У првом реду, то је начин реализације и пропусни опсег Т/Н кола. Уколико је вредност $|x_g|$ мала у односу на напонски опсег улаза у Т/Н коло, његов фреквентни опсег се може повећати два до три реда величине тако да и ретки догађаји, каква је вероватноћа $p(|x| < |x_g|)$, могу да имају прихватљиву учесталост. Прелиминарне анализе су показале да учесталост таквог догађаја може да иде, са доступном технологијом Т/Н кола, до 20 MHz. Исцрпна анализа тог проблема превазилази оквире овог рада. Предлог, изнет у раду, је намењен научној и стручној јавности у области мерења и метрологије, јер за више од реда величине подиже фреквентни опсег СДММ и обезбеђује аналогну (бесконачно велику) резолуцију дитерског сигнала.

V. ЗАКЉУЧАК

У раду је изнет предлог идеје новог генератора аналогног дискретног униформног шума. Дата је његова идејна хардверска шема, описан је принцип његовог рада и наведене су његове главне предности: бесконачно велика резолуција и бар за ред величине већи фреквентни пропусни опсег. Озбиљан инжењерски напор и труд мора бити уложен да би се од њега направило решење које и по осталим параметрима мора да буде на нивоу стандардног решења, а то је, у првом реду, тачност.

ЛИТЕРАТУРА

[1] V. Vujicic ***et al.***, “Low frequency stochastic true RMS instrument", *IEEE Trans. Instrum. Meas.*, vol. 48, no. 2, pp. 467-470, Apr. 1999.

[2] V. Vujicic *et al*., "Measurement in a point versus measurement over an interval", *Proc. of the IMEKO XIX World Congress*, pp. 1128-1132, Sep. 2009.

[3] M. Urekar *et al.*, “Accuracy Improvement of the Stochastic Digital Electrical Energy Meter,” *Measurement*, vol. 98, pp. 139-150, Feb. 2017.

[4] D. Pejic *et al*., “Stochastic Digital DFT Processor and Its Application to Measurement of Reactive Power and Energy,” *Measurement*, vol. 124, pp. 494–504, Aug. 2018.

[5] P. Sovilj *et al*., “Synergy and completeness of simple A/D conversion and simple signal processing,” *J. Phys.: Conf. Ser.,* vol. 127, Nov. 2019.

[6] V. Vujicic *et al.,* ”Mathematical Basis of the Stochastic Digital Measurement Method”, *Proc. of the 2nd Int. Conf. on Stochastic Processes and Algebraic Structures*, pp. 55-56, Oct. 2019.

[7] V. Vujicic *et al*., "Stochastic Measurement of Wind Power using a Two-bit A/D Converter", Measurement, vol. 152, paper no. 107184, Feb. 2020.

[8] M. Urekar, Contribution to the optimization of digital measurements performance (in Serbian), Ph.D. dissertation, Faculty of Technical Sceinces, University of Novi Sad.

[9] V. Vujicic, D. Pejic and A. Radonjic, "A brief overview of stochastic instruments for measuring flows of electrical power and energy", *FU Elec. Energ.*, vol. 32, no. 3, pp. 439-448, Sept. 2019.

[10] V. Vujicic *et al*., “Measurement in Fourier domain - a Natural Method of Big Data Volume Reduction", *Proc. of the 6th Int. Conf. on Electrical, Electronic and Computing Engineering (IcETRAN 2019)*, pp. 471-474, June 2019.

[11] Analog Devices Inc., AD5791 Series Digital to Analog Converters, https://www.analog.com/en/products/ad5791.html#

ABSTRACT

The paper proposes the idea of an analog discrete uniform noise generator for application in the stochastic digital measurement method. The proposed solution, unlike the standard one, which uses a random number generator and a precise and accurate digital-to-analog converter (DAK), overcomes the problem of limited DAK resolution and its narrow bandwidth.

Preliminary design of analog discrete uniform noise generator

Bojan Vujicic, Dragan Pejic, Aleksandar Radonjic and Vladimir Vujicic

Primer daljinskog merenja sinusoidalnih signala instrumentom VMP20

Jovan Ničković, Jelena Đorđević Kozarov, Radoje Jevtić i Atila Juhas

Apstrakt— **Monofazni analizator snage VMP20 koristi se za realizaciju nastave u Laboratoriji za elektroenergetiku u ETŠ „Nikola Tesla“ u Nišu. U radu je prikazana realizacija laboratorijske vežbe, koja se izvodi u okviru predmeta Merenja u elektroenergetici. Dobijeni rezultati su prikazani tabelarno. U radu se, takođe, opisuje i daljinsko merenje sinusoidalnih signala u ED mreži korišćenjem VMP20 monofaznog analizatora snage. Demonstracija on-line prikaza mrežnog napona i frekvencije realizovana je na samoj konferenciji ETRAN 2021.**

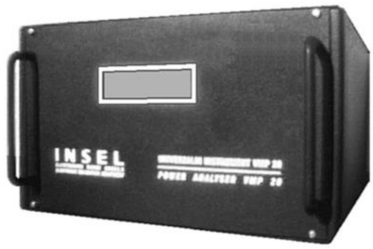
Ključne reči— **Monofazni analizator snage VMP20; Daljinsko merenje; Merenje sinusoidalnih signala u ED mreži.**

I. Uvod

RAZVOJ metoda za digitalno stohastičko merenje stacionarnih signala opisan je u literaturi [1-4]. Opisana stohastička digitalna merna metoda (SDMM) karakteriše se izuzetno jednostavnim, robusnim i pouzdanim hardverom visoke tačnosti. SDMM omogućava potpuno paralelna merenja, kao i paralelnu obradu mernih podataka. S obzirom da SDMM koristi dvobitne fleš A/D konvertore, propusni opseg je vrlo širok, dok se ključna operacija u obradi podataka MAC (Multiply and Accumulate) izvodi u jednom taktu zahvaljujući jednostavnom hardveru. Posledica toga je velika brzina obrade izmerenih podataka u vremenskom i frekventnom domenu, što omogućava i vrlo efikasan nadzor nad elektrodistributivnom (ED) mrežom i tehnološkim procesima [5].

II. Instrument VMP20

Instrument VMP20, monofazni analizator snage, projektovan je na Katedri za električna merenja, pri Fakultetu tehničkih nauka Univerziteta u Novom Sadu, i proizveden 1996. godine. Baziran je na patentu [6] i namenjen je merenjima u ED mreži na mrežnoj frekvenciji od 50 Hz. Direktno meri napon, struju i aktivnu snagu (direktan prikaz na displeju), a indirektno meri frekvenciju, impedansu, reaktivnu snagu, prividnu snagu, aktivnu i reaktivnu energiju (dobijaju se uz pomoć softvera na računaru). Instrument ima dva „plivajuća“ kanala - strujni i naponski. Oba kanala su baždarena sinusnim veličinama, na frekvenciji od 50 Hz. Tačnost merenja napona i struje je 0,5% FS, aktivne snage je 1% FS i frekvencije je 0,02% FS. Dometi su 400 V na naponskom, odnosno 5 A na strujnom kanalu. Na Sl. 1 prikazan je instrument VMP20.



Sl. 1. Instrument VMP20 – monofazni analizator snage

Na displeju instrument pokazuje napon u voltima, struju u amperima, aktivnu snagu u vatima i apsolutnu vrednost faktora snage sa najviše 3 značajne cifre. Klasa faktora snage je 2% od punog opsega. Instrument sračunava i prikazuje na displeju apsolutnu vrednost faktora snage.

Sve četiri veličine koje se mere (napon, struja, aktivna snaga i frekvencija) su predstavljene svojim srednjim vrednostima u 100 perioda mrežne učestanosti, dakle u dvosekundnom vremenskom intervalu. Merenje je sinhronizovano sa učestanošću mrežnog napona.

A. Softver VMPCalc

Pored instrumenta realizovan je i softver za podršku VMPCalc, koji se koristi za dodatnu obradu podataka dobijenih merenjem instrumentom VMP20. Monofazni instrument VMP20 u jednoj fazi direktno meri, koristeći SDMM, efektivne vrednosti napona i struje, aktivnu snagu i frekvenciju mreţnog napona. Izmerene vrednosti se pomoću povezanog PC računara pohranjuju u bazu podataka svake sekunde.

Program VMPCalc obrađuje snimljene podatke i izračunava prividnu snagu, faktor snage, moduo impedanse, Fryze-ovu reaktivnu snagu. Sve podatke, i izmerene i izračunate, izvozi u Excel tabele sa grafikonima.

Program VMPCalc, osim izvoza izmerenih i izračunatih podataka u tabele, izračunava srednje vrednosti, minimum, maksimum i standardnu devijaciju svih direkno izmerenih i izračunatih veličina, kao i maksimalne petnaestominutne prosečne vrednosti aktivne snage (vršna snaga), reaktivne

Jovan Ničković – ETŠ “Nikola Tesla”, Aleksandra Medvedeva 18, 18000 Niš, Srbija (e-mail: jovan@etstesla.ni.ac.rs).
Jelena Đorđević Kozarov – Elektronski fakultet, Univerzitet u Nišu, Aleksandra Medvedeva 14, 18000 Niš, Srbija (e-mail: kozarov@elfak.ni.ac.rs).
Radoje Jevtić – ETŠ “Nikola Tesla”, Aleksandra Medvedeva 18, 18000 Niš, Srbija (e-mail: milan.jvtc@gmail.com).
Atila Juhas – Fakultet tehničkih nauka, Univerzitet u Novom Sadu, Trg Dositeja Obradovića 6, 21000 Novi Sad, Srbija, (e-mail: juhiatti@gmail.com).

snage, prividne snage i struje. Sve ove podatke snima kao kratak izveštaj obrade.

U bazi podataka je moguće uzeti proizvoljni vremenski podinterval, u kojem je moguće izračunati aktivnu energiju *EP*, reaktivnu energiju *EQ* i prividnu energiju *ES*.

S obzirom da se u bazu podataka svake sekunde upisuju aktuelni mereni podaci, zapisi o snagama *i*-te sekunde (*P*i, *Q*i i *S*i) brojčano su jednaki energijama za taj vremenski interval od jedne sekunde (*EP*i, *EQ*i, *ES*i). Enegije se mogu izračunati kao prosta suma svih pojedinačnih energija unutar izabranog vremenskog intervala. Ovaj način je pogodan za računanje energija pomoću Excel tabela.

U kratkom izveštaju koriste se proračuni dobijeni programom VMPCalc, koji računaju srednje vrednosti aktivne snage *P*sr, reaktivne snage *Q*sr i prividne snage *S*sr u vremenskom intervalu *T*. Energije se u ovom slučaju dobijaju prostim množenjem srednjih vrednosti sa dužinom vremenskog intervala.

B. *Korisnički interfejs programa VMPCalc*

Na Sl. 2 prikazan je korisnički interfejs programa VMPCalc. Ispod naslovne linije nalaze se komande za upravljanje programom. Komandom *Otvori bazu* otvara se baza podataka i vrši se tabelarni prikaz cele baze u prozoru koji se nalazi ispod dugmeta. Komandama *Pocetno vreme* i *Krajnje vreme* podešavaju se početno i krajnje vreme intervala obrade. Komandom *Podesi interval*, nakon podešavanja granica intervala obrade, vrši se priprema podataka za obradu. Nakon što je interval podešen, komanda menja naziv u *Resetuj interval*, čijim se odabirom ponovo priprema celokupna baza podataka za prikaz. Komandom *Proračun* vrši se proračun i prikaz podataka, kao i generisanje izveštaja u desnom oknu, koje je namenjeno za prikaz izveštaja. Komandom *Izvezi u Excel* odabrani interval podataka iz baze izvozi se u Excel tabele. Komandom *Sacuvaj izvestaj* snima se izveštaj kao tekstualna datoteka u tekućem folderu programa. Komandom *Obrisi izvestaj* briše se sadržaj prozora izveštaja.

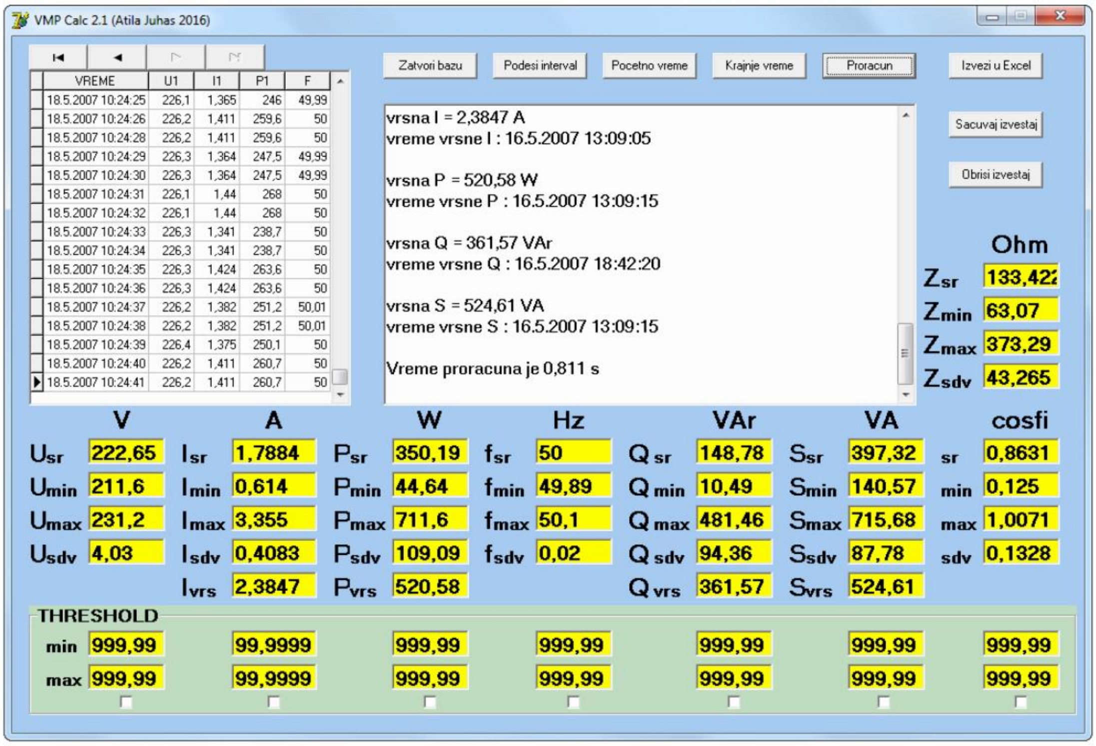
Ispod okna za prikaz podataka iz tabele i izveštaja, nalaze se displeji za prikaz izračunatih veličina organizovanih u kolone.

Ispod displeja za prikaz izračunatih podataka, nalazi se uokvireno polje za podešavanje *threshold*-a. Odabirom određenog *threshold*-a, iz obrade se izbacuju podaci koji su ispod minimuma ili iznad maksimuma date veličine. Nakon podešavanja *threshold*-a potrebno je ponovo aktivirati komandu *Podesi interval*.

Naravno, da bi funkcionisala opcija *Izvezi u Excel* na računaru mora biti instaliran Microsoft Office programski paket sa aplikacijom Microsoft Excel.

III. PRIMENA INSTRUMENTA VMP20 U NASTAVI

Monofazni mrežni analizator snage VMP20 koristi se za realizaciju nastave u Laboratoriji za elektroenergetiku elektrotehničke škole „Nikola Tesla“ u Nišu od školske 2019/2020. godine. Instrument se koristi za realizaciju laboratorijske vežbe, koja se izvodi u okviru predmeta Merenja u elektroenergetici.



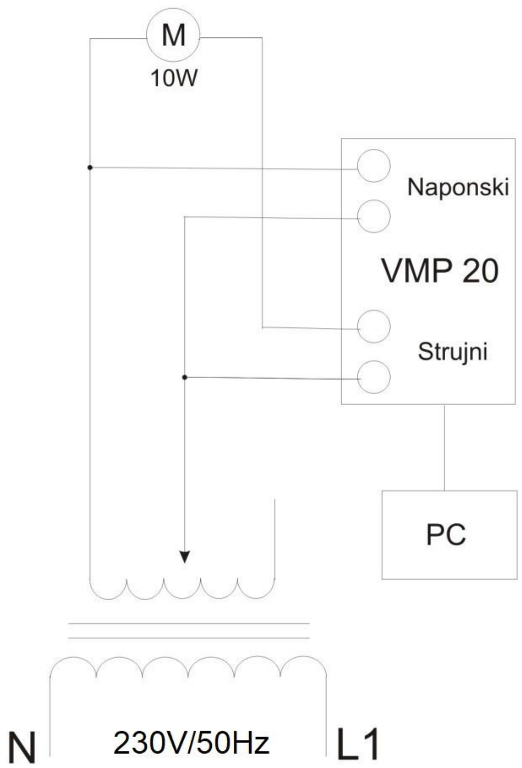
Sl. 2. Izgled VMPCalc po završetku obrade podataka.

TABELA I
REZULTATI MERENJA INSTRUMENTOM VMP20

U [V]	0	24,73	51,24	73,79	101,4	125,1	151,6	175,5	199,4	222,9	239,3
I [A]	0	0,023	0,041	0,059	0,095	0,132	0,173	0,171	0,198	0,223	0,248
P [W]	0	0,425	1,674	3,352	6,719	10,56	17,19	19,77	24,8	30,78	35,87
cos ϕ	0	0,76	0,802	0,768	0,7	0,641	0,654	0,659	0,632	0,618	0,604
f [Hz]	0	49,99	50	50	50	50,01	50	50,02	50	50	49,99

Zadatak u nastavnoj vežbi jeste da se izmere struja, snaga i frekvencija pri promeni napona napajanja. Napon se menja pomoću jednofazni regulacioni autotransformator u opsegu od 0 V do 250 V. Kao potrošač korišćen je monofazni motor sledećih karakteristika: FAN za 0,23 A, 220 V i 50 Hz. Ostale karakteristike motora su: 1350 RPN, snage 10 W.

Šema veza laboratorijske vežbe data je na Sl. 3, dok je na Sl. 4 prikazan instrument VMP20 u Laboratoriji.



Sl. 3. Šema veza laboratorijske vežbe sa monofaznim analizatorom snage VMP 20.

IV. REZULTATI MERENJA

Rezultati izvršenih merenja prikazani su u Tabeli I.

Za snimanje izmerenih podataka dobijenih pomoću uređaja VMP20 na PC računar upotrebljen je program VMPLOG. Program je predviđen da radi na standardnom PC računaru pod operativnim sistemom Microsoft Windows XP ili novijim. Aplikacija koristi Firebird tip baze podataka sa integrisanim klijent/server rešenjem.

Uz pomoć pratećih softvera moguće je i daljinsko merenje sinusoidalnih signala u ED mreži korišćenjem VMP20 monofaznog analizatora snage. Demonstracija on-line prikaza mrežnog napona i frekvencije realizovana je na samoj konferenciji ETRAN 2021.

Sl. 4. Instrument VMP20 u Laboratoriji.

V. ZAKLJUČAK

U radu je prikazana primena monofaznog analizatora snage VMP20 u realizaciji nastave u Laboratoriji za elektroenergetiku u ETŠ „Nikola Tesla“ u Nišu. U radu je prikazana realizacija laboratorijske vežbe, koja se izvodi u okviru predmeta Merenja u elektroenergetici. Isto tako, realizovana je demonstracija on-line prikaza merenja mrežnog napona i frekvencije na samoj konferenciji ETRAN 2021. S obzirom da se u protekle dve godine javila potreba za izvođenjem nastave na daljinu, uključivanje ovog instrumenta u nastavu se pokazalo kao veoma svrsishodno.

ZAHVALNICA

Ovaj rad je podržan od strane Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije.

LITERATURA

[1] V. Vujičić, S. Milovančev, M. Pešaljević, D. Pejić, and I. Župunski, “Low frequency stochastic true RMS instrument,” *IEEE Trans. Instrum. & Measurement,* vol. *48*, no. *2*, pp. *467–470*, Apr. 1999.

[2] V. Vujičić, “Generalized low frequency stochastic true RMS instrument,” *IEEE Trans. Instrum. & Measurement*, vol. *50*, no. *5*, pp. *1089–1092*, Oct. 2001.

[3] V.V. Vujicic, I. Župunski, Z. Mitrovic and M.A. Sokola, “Measurement in a point versus measurement over an interval,” Proc. of the IMEKO XIX World Congress; Lisbon, Portugal. no. 480, pp. 1128-1132, Sep. 6−11, 2009.

[4] V. Pjevalica and V. Vujičić, “Further generalization of the low-frequency true-RMS instrument,” *IEEE Trans. Instrum. & Measurement*, vol. *59*, no. *3*, pp. *736–744*, Mar. 2010.

[5] A. Juhas, P. Sovilja, V. Vujičić, “Softverska i hardverska podrška trofaznim mjerenjima u ED mreži,” Naučno-stručni simpozijum Energetska efikasnost – ENEF 2017, str. 84-89, Banja Luka, 3-4. Novembar 2017.

[6] Vladimir Vujičić, Slobodan Milovančev, „Patent P-742/95, Digitalni instrument za merenje proizvoda dva analogna periodična signala“, Zavod za intelektualnu svojinu Beograd, 1995.

ABSTRACT

The single-phase power analyzer VMP20 is used for the realization of teaching exercises in the Laboratory for Electric Power Engineering in ETS "Nikola Tesla" in Niš. The paper presents the realization of a laboratory exercise that is carried out within the course of the Measurements in Power Engineering. The obtained results are shown in the table. The paper also describes the remote measurement of sinusoidal signals in the ED network using a VMP20 single-phase power analyzer. Demonstration of on-line measurement of voltage and frequency is realized at the Conference ETRAN 2021.

Example of remote measurement of sinusoidal signals with VMP20 instrument

Jovan Ničković, Jelena Đorđević Kozarov, Radoje Jevtić and Atila Juhas

Indeks autora/Author index

CIP - Каталогизација у публикацији
Народна библиотека Србије, Београд

621.3(082)(0.034.2)
534(082)(0.034.2)
004(082)(0.034.2)
681.5(082)(0.034.2)
621.039(082)(0.034.2)
66.017(082)(0.034.2)
57+61(048)(0.034.2)
006.91(082)(0.034.2)

INTERNATIONAL Conference on Electrical, Electronic and Computing Engineering (8 ; 2021 ; Bijeljina)
Зборник радова [Електронски извор] / 65. годишња конференција за електронику, телекомуникације, рачунарство, аутоматику и нуклеарну технику ЕТРАН 2021 и 8. интернационална конференција за електротехнику, електронику и рачунарство ИцЕТРАН 2021, Етно село Станишићи [Бијељина], 08-10.09.2021. године = Proceedings / 8th International Conference on Electrical, Electronic and Computing Engineering IcETRAN 2021 and 65th National Conference on Electronics, Telecommunication, Computing, Automatic Control and Nuclear Engineering ETRAN 2021 ; [главни уредник Слободан Вукосавић = editor in charge Slobodan Vukosavić]. - Београд : Друштво за ЕТРАН : Академска мисао = Belgrade : ETRAN Society : Academic Mind, 2020 (Београд : Академска мисао). - 1 електронски оптички диск (CD-ROM) ; 12 cm

Системски захтеви: Нису наведени. - Насл. са насловне стране документа. - Радови на срп. и енгл. језику. - Тираж 200. - Библиографија уз сваки рад. - Abstracts.

ISBN 978-86-7466-894-8 (АМ)

1. Друштво за електронику, телекомуникације, рачунарство, аутоматику и нуклеарну технику (Београд). Конференција (65 ; 2021 ; Бијељина)

а) Електротехника - Зборници b) Акустика - Зборници c) Рачунарска технологија - Зборници d) Системи аутоматског управљања - Зборници e) Нуклеарна техника - Зборници f) Технички материјали - Зборници g) Биомедицина - Зборници h) Метрологија - Зборници

COBISS.SR-ID 51195145